

MIT LIBRARIES

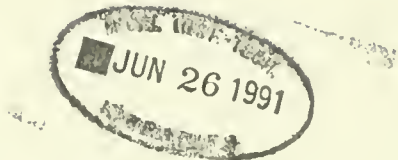


3 9080 00686241 8



HD28
.M414
no.
3304-
91

DEWEY



WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

An Overview of Modern Financial Economics

by

Chi-fu Huang
Massachusetts Institute of Technology

WP# 3304-91-EFA

June 1991

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

An Overview of Modern Financial Economics

by

Chi-fu Huang
Massachusetts Institute of Technology

WP# 3304-91-EFA

June 1991



An Overview of Modern Financial Economics¹

Chi-fu Huang

1 Introduction

The seemingly simple observation made by Modigliani and Miller (1958) that two financial packages having identical payoffs in all states of the world must sell for the same price marked the beginning of the modern financial economics, where formal economic reasonings replaces heuristics. In this paper we will briefly overview three key subject areas in the modern financial economics relating to the capital market theory. We begin by introducing the mean-variance portfolio theory of Markowitz (1952) and Tobin (1958) and its implication on equilibrium expected asset returns, or the so-called *Capital Asset Pricing Model*, due to Black (1972), Lintner (1965), Mossin (1965), and Sharpe (1964). We then turn our attention to the generalizations of the above theories to dynamic economies as pioneered by Merton (1971, 1973a) and later extended by Breeden (1979) and Cox, Ingersoll, and Ross (1985). The models that work out equilibrium implications on expected asset returns are termed the *Intertemporal Capital Asset Pricing Models*.

The path-breaking work that has spurred more than a decade of intensive research and numerous innovations in financial engineering is the third and final subject on our agenda. It is the valuation theory of options originated by Black and Scholes (1973) and Merton (1973b). The theory now has evolved into a general theory of contingent claims valuation using probabilistic methods due mainly to the work of Cox and Ross (1976), Harrison and Kreps (1979), and Harrison and Pliska (1981). As an outgrowth of this general theory, we now also have a dynamic portfolio theory and a dynamic equilibrium theory using probabilistic methods, the former due to Cox and Huang (1986, 1989), He and Pearson (1989), and Pagès (1989) and the latter due to Huang (1985a, 1985b), Duffie and Huang (1985), Duffie (1986), and Huang (1987).

More detailed discussions of these three areas and some applications of the key ideas will appear in four short articles following this overview.

2 Static Portfolio Theory and the Capital Asset Pricing Model

The static portfolio theory concerns with finding the portfolio of assets that has a given level of expected rate of return and has a minimum risk exposure measured by the variance of the return on a portfolio. A portfolio having this property will be said to be a mean-variance efficient portfolio.

Suppose to begin that there is no riskless asset. Then one can show that the set of mean-variance efficient portfolios, or the *portfolio frontier*, is a hyperbola in the mean-standard deviation space; see the hyperbola in Figure 1. Among all the portfolios on the portfolio frontier is a portfolio

¹This is a collection of five articles whose Japanese translations have appeared in five issues of the *Communications of Operations Research of Japan* from 1990 to 1991 with a co-author Tadashi Uratani

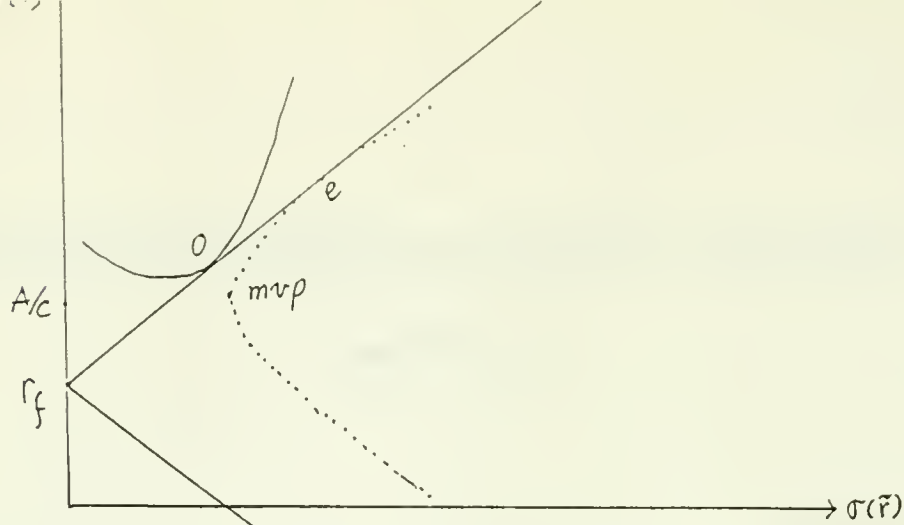


Figure 1: A Portfolio Frontier

with the minimum variance, termed henceforth as the *minimum variance portfolio*. It is known that the complete portfolio frontier is “generated” by any two distinct portfolios on the frontier in that any portfolio on the frontier can be written as a linear combination of two given frontier portfolios. This is called the *two fund separation* property. The implication is that if investors want to hold frontier portfolios, then there just needs to have any two frontier portfolios available as mutual funds to satisfy all investors since any other frontier portfolio is just a portfolio of these two mutual funds. Conversely, one can demonstrate that the portfolio frontier is a linear space in that any linear combination of frontier portfolios is itself a frontier portfolio.

Next one can show that for any portfolio p on the portfolio frontier that is not the minimum variance portfolio, there exists another frontier portfolio, denoted by $zc(p)$, whose rate of return has a zero covariance with respect to that on p . Geometrically, $zc(p)$, the *zero covariance portfolio* of p , is easily found. One just draws a line tangent to the portfolio frontier at p . The intercept of this line on the expected rate of return axis gives the expected rate of return of $zc(p)$; see Figure 1. Now note the striking mathematical result: fix a frontier portfolio p and its zero covariance portfolio $zc(p)$. Let q be any portfolio not necessarily on the frontier. Then we have

$$E[\tilde{r}_q] - E[\tilde{r}_{zc(p)}] = \beta_{qp}(E[\tilde{r}_p] - E[\tilde{r}_{zc(p)}]), \quad (1)$$

where $\tilde{r}_i$ denotes the rate of return on portfolio i , $E[\cdot]$ is the expectation operator, and $\beta_{qp} \equiv \frac{\text{Cov}(\tilde{r}_q, \tilde{r}_p)}{\text{Var}(\tilde{r}_p)}$ with Cov and Var denoting covariance and variance, respectively.

Equation (1) says that, for a fixed frontier portfolio p which is not the minimum variance portfolio, there exists a linear relation between the expected rate of return of any portfolio q and β_{qp} . This is a mathematical result and is void of economic content. The contribution of the *Capital Asset Pricing Model* lies in the identification of a frontier portfolio using economic arguments and thus produces a linear restriction among the expected rate of returns on assets. In this model, hypotheses either about asset returns or about individuals’ preferences are made so that individuals will optimally choose to hold frontier portfolios. In equilibrium, supply must be equal to demand for the markets to clear. Thus the market portfolio is a convex combination of individuals’ optimal portfolios and is itself a frontier portfolio. Certain economic hypothesis about nonsatiation is then made to ensure that the market portfolio has an expected rate of return strictly higher than that on the minimum variance portfolio. It follows that (1) holds with p replaced by the market portfolio m and $zc(p)$ replaced by $zc(m)$, and the term $E[\tilde{r}_m] - E[\tilde{r}_{zc(m)}]$ is strictly positive. This is the *Zero-Beta Capital Asset Pricing Model* of Black (1972).

The case where there exists a riskless asset in zero net supply is similar. The portfolio frontier is composed of two half-lines emanating from r_f , where r_f is the riskless interest rate; see the two half-lines in Figure 1. We still have the two fund separation property. Here, the two separating mutual funds can always be chosen to be the riskless asset and the “tangent” portfolio e . Note that the tangent portfolio e is a portfolio composed only of risky assets and is on the portfolio frontier constructed solely using risky assets. Now suppose that individuals optimally choose to hold frontier portfolios (of all assets). Since the riskless asset is in zero supply, market clearing condition implies that in equilibrium e must be the market portfolio. Hence the expected rate of return on $zc(m)$ is equal to r_f by the geometry of the portfolio frontier. It then follows from relation (1) that for any portfolio q , we have

$$E[\tilde{r}_q] - r_f = \beta_{qm}(E[\tilde{r}_m] - r_f), \quad (2)$$

where $\beta_{qm} \equiv \frac{\text{Cov}(\tilde{r}_q, \tilde{r}_m)}{\text{Var}(\tilde{r}_m)}$ and $\tilde{r}_m$ is the rate of return on the market portfolio.

Calling the term $E[\tilde{r}_q] - E[\tilde{r}_{zc(m)}]$ the risk premium on portfolio q , both the Capital Asset Pricing Model and the Zero-Beta Capital Asset Pricing Model predict that the risk premium on any asset is proportional to its “beta” with respect to the market portfolio. The higher the beta the larger the risk premium. The “beta’s” with respect to the market portfolio instead of the variances of returns are the measure of risk in equilibrium. This is the *Capital Asset Pricing Model* of Lintner (1964), Mossin (1965), and Sharpe (1964).

Both the Capital Asset Pricing Model and the Zero-Beta Capital Asset Pricing Model formalize the following economic intuition: Consider two assets, a and b , having the same expected dividends. The dividends of asset a is positively correlated with those of the market portfolio while the dividends of asset b is negatively correlated with those of the market portfolio. Asset b , as compared to asset a , provides a “hedge” against the aggregate fluctuation and thus individuals find it an attractive investment relative to asset a . In equilibrium all assets will be held and thus individuals holding asset a must be rewarded by a higher expected rate of return. The asset pricing models predict that the risk premiums are a linear increasing function of assets’ “betas” with respect to the market portfolio. The part of the randomness of the return on an asset that is uncorrelated with $\tilde{r}_m$ does not contribute to the asset’s risk premium as it can be diversified away.

3 Dynamic Portfolio Theory and the Intertemporal Capital Asset Pricing Model

The dynamic portfolio theory analyzes an individual’s problem of investing his wealth and withdrawing from his portfolio to consume to maximize his preferences for consumption over time. For analytic tractability, the theory has been developed in continuous time economies where asset prices follow a diffusion process.

Formally, consider an economy under uncertainty with a time horizon $[0, T]$. There are $N + 1$ securities traded indexed by $n = 0, 1, \dots, N$. Let $S_n(t)$ denote the price of security n at time t and let $S(t)$ be $(S_1(t), \dots, S_N(t))^T$, where ^{T} denotes “transpose.” Assume that S satisfies the stochastic differential equation:

$$dS(t) + f(Y(t), t)dt = I_S(t)\mu(Y(t), t)dt + I_S(t)\sigma(Y(t), t)dB(t), \quad (3)$$

where Y is an M -vector diffusion process satisfying the stochastic differential equation

$$dY(t) = \beta(Y(t), t)dt + g(Y(t), t)dB(t), \quad (4)$$

$f(Y(t), t)$ is an N -vector of dividend rates paid by securities $n = 1, 2, \dots, N$ at time t , $I_S(t)$ is an $N \times N$ diagonal matrix process with the n -th diagonal element being equal to $S_n(t)$, μ is an N -vector processes, β is an M -vector process, σ is an $N \times K$ -matrix process, g is an $M \times K$ matrix process, and B is a K -dimensional standard Brownian motion. To aid our future interpretations, let $M + 1 < N \leq K$ and assume that $\sigma(t)$ and $g(t)$ are of full rank for all t . We will call Y the vector of “state variables,” and $\mu(t)$ and $\sigma(t)\sigma(t)^\top$ are the vector of “instantaneous” expected rate of returns and the “instantaneous” covariance matrix, respectively, on assets $n = 1, 2, \dots, N$ at time t . The processes S and Y form a diffusion process. Security 0 is locally riskless with a price process growing at a rate $r(Y(t), t)$ at time t . Henceforth, we will call securities $n = 1, 2, \dots, N$ risky assets and security $n = 0$ the riskless asset.

A portfolio policy A is an N -vector process that prescribes the proportion of wealth invested in the risky assets and a consumption policy C is a positive process that describes the consumption rates over time in the single consumption good. For a given pair (A, C) , the budget feasible wealth process W must satisfy the following dynamics:

$$dW(t) = [W(t)(r(t) + A(t)^\top(\mu(t) - r(t))) - C(t)]dt + W(t)A(t)^\top\sigma(t)dB(t). \quad (5)$$

The optimal consumption and portfolio problem of an individual is finding a pair (A, C) to maximize

$$E \left[\int_0^T u(C(t), Y(t), t) dt \right]$$

subject to the budget constraint (5) and price dynamics (3) and (4), where $u(C(t), Y(t), t)$ is an individual's felicity function for time t consumption. Assume there exists a solution to the above problem and let $J(W(t), Y(t), t)$ be the optimal value function for this problem at time t given that wealth and state variables at that time are $W(t)$ and $Y(t)$, respectively. Bellman's optimality principle implies that J must satisfy the following Bellman equation:

$$\begin{aligned} 0 = \max_{A, C} \left\{ \frac{1}{2} J_{WW} W^2 A^\top \sigma \sigma^\top A + W A^\top \sigma g^\top J_{WY} \right. \\ \left. + J_W [W(r + A^\top(\mu - r)) - C] + u(C, Y, t) \right\} \\ + J_Y^\top \beta + \frac{1}{2} \text{tr}(J_{YY} g g^\top) + J_t, \end{aligned} \quad (6)$$

where subscripts denote partial derivatives and tr denotes “trace.”

Performing maximization gives

$$A = -\frac{J_W}{J_{WW}W}(\sigma\sigma^\top)^{-1}(\mu - r) - (\sigma\sigma^\top)^{-1}\sigma g^\top \frac{J_{WY}}{J_{WW}W}, \quad (7)$$

$$u_c \leq J_W, \quad (8)$$

where u_c denotes partial derivative of u with respect to c . Note that the inequality of (8) holds as an equality if consumption is strictly positive.

Relation (7) exhausts the implications of the dynamic portfolio theory. One can show that $(\sigma\sigma^\top)^{-1}(\mu - r)$ is proportional to the portfolio weights on the risky assets of an “instantaneous” mean-variance frontier portfolio, and the i -th row of $(\sigma\sigma^\top)^{-1}\sigma g^\top$ is proportional to the portfolio weights on the risky assets of the portfolio mostly highly correlated with the i -th state variable. We immediately have the generalization of the two-fund separation result in the static setting: all the optimal portfolios are linear combinations of a mean-variance frontier portfolio, M portfolios most highly correlated with the state variables, and the riskless asset. We thus have $M + 2$ fund separation. The hypothesis that $M + 1 < N$ made earlier is for this “mutual fund separation” to be meaningful, or else, the number of mutual funds would not be smaller than the number of assets traded.

Given that there exists an equilibrium in our dynamic economy in which there are state variables Y so that prices for securities follow the processes we posited, equilibrium relation among expected rates of return of assets can be derived as follows. For simplicity, assume that there exists a single representative individual in the economy. Then the market clearing condition implies that (7) must hold for the portfolio weights of the market portfolio while the wealth of the representative individual must be the aggregate wealth. Using these observations, Itô’s lemma and (7) imply

$$\begin{aligned} (\mu(t) - r(t))dt &= -\frac{J_W W(t) W(t)}{J_W(t)} \text{Cov}_t((dW(t) + C(t)dt)/W(t), (dS(t) + f(t)dt)/S(t)) \\ &\quad - \frac{J_W Y(t)}{J_W(t)} \text{Cov}(dY(t)^\top, (dS(t) + f(t)dt)/S(t)) \\ &= -\text{Cov}_t(dJ_W(t)/J_W(t), (dS(t) + f(t)dt)/S(t)), \end{aligned} \quad (9)$$

where Cov_t denotes the covariance operator conditional on the information at time t . Note that $(dS + fdt)/S$ is the vector of instantaneous rates of return on risky assets and $(dW + Cdt)/W$ is the rate of return on the aggregate wealth/market portfolio.

The second equality of (9) underlies much of the intertemporal asset pricing models. The equilibrium instantaneous expected rate of return on an asset depends on how its rate of return is correlated with the changes in the representative individual’s marginal utility for wealth as captured by dJ_W . If its rate of return is positively correlated with dJ_W , then its equilibrium risk premium will be negative since it provides a hedge by offering high returns when the representative individual needs it the most.

Now we will derive the multi-factor *Intertemporal Capital Asset Pricing Model* of Merton (1973a). Let X be a vector of value processes of M portfolio most highly correlated with the state variables Y and let $\mu_x(t)$ denote the M -vector of instantaneous expected rates of return per unit time on X . By manipulating the first equality of (9), we get

$$\mu(t) - r(t) = \beta_{S, WX} \begin{pmatrix} \mu_m(t) - r(t) \\ \mu_x(t) - r(t) \end{pmatrix}, \quad (10)$$

where

$$\beta_{S, WX} = V_{S, WX}(t) V_{WX, WX}^{-1}(t),$$

$V_{S, WX}(t)$ and $V_{WX, WX}(t)$ are the covariance matrices of the instantaneous changes at time t in S and $(W, X^\top)$ and in $(W, X^\top)$ and $(W, X^\top)$, respectively, and $\mu_m(t)$ denotes the instantaneous expected rate of return at time t on the market portfolio. That is, in equilibrium there exists a linear relation between the instantaneous risk premiums on risky assets and their $M + 1$ characteristics

as captured by the rows of $\beta_{S,WX}$ or the assets' beta's with respect to the market portfolio and with respect to the M portfolios most highly correlated with the state variables. The weights on these characteristics are the risk premiums on the market portfolio and on the M portfolios most highly correlated with the state variables. This is the *Intertemporal Capital Asset Pricing Model* of Merton (1973a).

With two more assumptions, we can simplify (10) further. Suppose now that the felicity function u is independent of the state variables and in equilibrium $C(t) > 0$ for all t . Using (8) with the inequality replaced by an equality, (9) becomes

$$(\mu(t) - r(t))dt = -\text{Cov}_t(du_c(t)/u_c(t), (dS(t) + f(t)dt)/S(t)). \quad (11)$$

Itô's lemma then implies that

$$(\mu(t) - r(t))dt = -\frac{u_{cc}(t)C(t)}{u_c(t)}\text{Cov}_t(dC(t)/C(t), (dS(t) + f(t)dt)/S(t)),$$

where we note that, by the market clearing condition, $C(t)$ is the time t aggregate consumption. Letting S_c denote the price process of the portfolio that does not pay dividend and whose rates of return are mostly highly correlated with the instantaneous rate of changes on the aggregate consumption, we have

$$\mu(t) - r(t) = \frac{\beta_{S_c}^\top}{\beta_{cc}}(\mu_c(t) - r(t)), \quad (12)$$

where $\mu_c(t)$ is the instantaneous expected rate of return at time t of S_c ,

$$\begin{aligned} \beta_{S_c} &\equiv \frac{\text{Cov}_t(dC(t)/C(t), (dS(t) + f(t)dt)/S(t))}{\text{Var}_t(dC(t)/C(t))}, \\ \beta_{cc} &\equiv \frac{\text{Cov}_t(dC(t)/C(t), dS_c(t)/S_c(t))}{\text{Var}_t(dC(t)/C(t))}, \end{aligned}$$

and Var_t denotes the variance operator conditional on the information at time t . This is the *Intertemporal Consumption Capital Asset Pricing Model* of Breeden (1979) and we call the elements of β_{S_c}/β_{cc} the *consumption beta's* of the risky assets. This relation is similar to the Capital Asset Pricing Model except that, in place of the market portfolio, the aggregate consumption plays a pivotal role in asset pricing and the riskiness of an asset is completely determined by its consumption beta.

4 Valuation of Contingent Securities

The observation made by Modigliani and Miller (1958) that two financial packages having identical payoffs in all states of nature must sell for the same price underlies the modern theory of valuation of contingent securities as pioneered by Black and Scholes (1973) and Merton (1973b). Although this literature has mostly been developed using continuous-time stochastic processes, much of the intuition, however, can be understood in a discrete-time discrete-state setting to follow.

Consider an economy with three dates, $t = 0, 1, 2$, and three securities. These three securities will be called *primitive securities*. The price processes of the primitive securities are depicted in

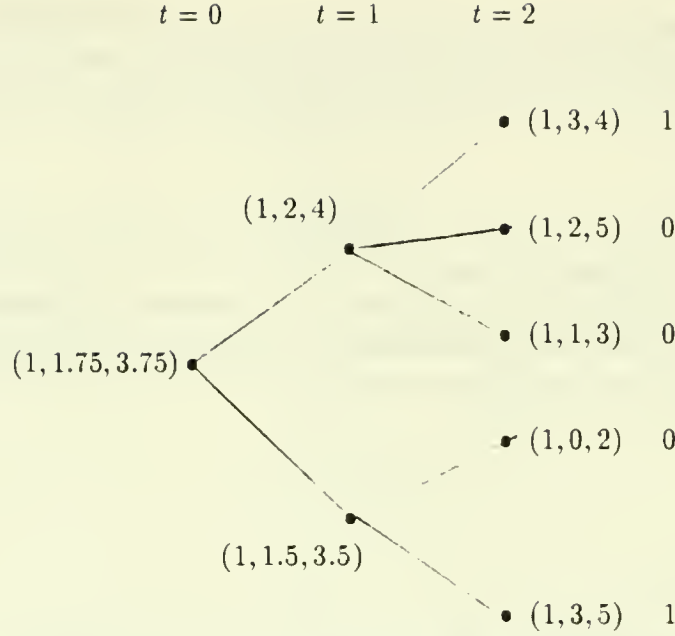


Figure 2: Price Processes Event Tree

Figure 2 in the parenthesis along the event tree. For example, the two sets of numbers in the parentheses at time 1 are the two scenarios of the primitive securities prices at time 1. Given that the realized prices at time 1 are those at the upper node, there are three possibilities at time 2 represented by the three branches succeeding the upper node. The n -th number of each parenthesis is the price of the n -th primitive security. Note that the first primitive security has a unit price throughout. Here we will not be specific about the probabilities of the branches on the tree except noting that each branch is of a strictly positive probability. Thus the implications we are going to derive will be independent of the probabilities.

We are interested in the valuation at time 0 of a security whose payoffs depend on, say, security 2. This security will be called a *contingent security* or a *derivative security*. Without loss of generality, we consider a European call option on the second primitive security that matures at time 2 with an exercise price 2. The numbers to the right of the parentheses at time 2 are the payoffs of this call option at that time. It is easily verified that the payoffs of this call option does not lie in the linear span of the time 2 prices of the three primitive securities. Thus one cannot construct a portfolio of the three primitive securities at time 0 and hold it to time 2 to have the payoffs of the call option. If such a portfolio existed, then the value of the call option would just be the value of this portfolio or else an arbitrage opportunity could be created by selling the call option and using the proceeds to buy this portfolio or vice versa.

The modern theory of the valuation of contingent securities makes the following observation: one can make use of the trading opportunity at time 1 to create payoffs at time 2 that do not lie in the linear span of the time 2 prices of the three primitive securities. At time 1 at the upper node, we ask: does there exist a portfolio of the three securities so that the value of the portfolio at time 2 will be identical to those of the call option in all three subsequent nodes? Letting x, y, z denote the numbers of shares of the three primitive securities held at time 1 on the upper node,

mathematically, we want to find a solution to the following system of equations:

$$\begin{aligned}x + 3y + 4z &= 1, \\x + 2y + 5z &= 0, \\x + y + 3z &= 0.\end{aligned}\tag{13}$$

The unique solution is $(x, y, z) = (1/3, 2/3, -1/3)$. That is, if at time 1 at the upper node one holds a portfolio into time 2 composed of one-third of a share of the first primitive security, two-thirds of a share of the second primitive security, and minus one-third of a share of the third primitive security, then at time 2, one has the exact payoffs of the option at the nodes subsequent to the upper nodes at time 1. The value of this portfolio at time 1 at the upper node is $1/3$. Similarly, for the lower node at time 1, we solve

$$\begin{aligned}x + 0y + 2z &= 0, \\x + 3y + 5z &= 1.\end{aligned}\tag{14}$$

A solution is $(x, y, z) = (-1/3, 1/6, 1/6)$ and the value of this portfolio at time 1 at the lower node is $1/2$.

Now if one can find a portfolio of the three primitive securities at time 0 so that its time 1 value is $1/3$ at the upper node and $1/2$ at the lower node, then one is done. This is so because one can then rebalance the compositions of the portfolio according to the solution to either (13) or (14) depending on whether the realized time 1 prices of the three primitive securities are at the upper node or the lower node. There will be neither additional funds needed for nor funds withdrawn out of the portfolio. The time 0 value of the call option will just be the value of the portfolio constructed at time 0. The time 0 composition of this portfolio is a solution to the equations:

$$\begin{aligned}x + 2y + 4z &= 1/3, \\x + 1.5y + 3.5z &= 1/2.\end{aligned}\tag{15}$$

A solution is $(x, y, z) = (4/3, -1/6, -1/6)$ and one has a complete specification of a portfolio whose time 2 payoffs are identical to those of the European call option. The value of this portfolio at time 0 is $5/12$ and thus the value of the call option then must also be $5/12$ since the prices for two financial packages having identical payoffs must be the same to prevent arbitrage opportunities.

More generally, one can constructed a portfolio of the three securities so that the portfolio has a unit payoff at one node on the event tree and nothing otherwise. Denote the time 0 price of the portfolio that pays one unit of wealth at time $t = 1, 2$ at node i by π_{ti}^* , where we have ranked the nodes at each time from top to the bottom. One first shows that

$$\sum_i \pi_{ti}^* = 1, \quad t = 1, 2$$

and

$$\begin{aligned}\sum_{i=1}^3 \pi_{2i}^* &= \pi_{11}^*, \\ \sum_{i=4}^5 \pi_{2i}^* &= \pi_{12}^*,\end{aligned}$$

where we note that the first equality is implied by the fact that the first primitive security has unit price throughout. That is, these prices satisfy the definition of a probability and can thus be

interpreted to be the probabilities of the respectively nodes. One verifies quickly that under these probabilities, the price processes for the primitive securities follow a martingale. Moreover, the no arbitrage arguments made earlier necessitate that a contingent security that is represented by its payoffs only at time 2, c_{2i} at node i , must sell for

$$\sum_{i=1}^5 \pi_{2i}^* c_{2i}$$

at time 0, or sell for the expectation according to the π_{2i}^* 's of its payoffs. Similarly, the prices of this contingent security at time 1 at the upper and lower nodes are, respectively,

$$\sum_{i=1}^3 \frac{\pi_{2i}^*}{\pi_{11}^*} c_{2i}$$

and

$$\sum_{i=4}^5 \frac{\pi_{2i}^*}{\pi_{12}^*} c_{2i},$$

which are the expectation at time 1 of the c_{2i} 's conditional on whether the realized primitive security prices are at the upper node or at the lower node. Therefore, the prices over time of the contingent security is a martingale under the “probability” that assigns the probability π_{ti}^* to the i -th node at time t . This representation facilitates our computation of the prices for other contingent securities as we can simply calculate the conditional expectation of their payoffs according to the probability represented by the π_{ti}^* 's. This probability is termed a *martingale measure* by its connection to martingales.

In our setup, a martingale measure exists and is unique as for every node on the event tree, there exists a portfolio that has unit payoff at that node and nothing otherwise and there is a primitive security with unit price throughout. If there is no primitive security with unit price throughout, we take one of the securities to be the numeraire and normalize the prices. Then there exists a unique martingale measure for the normalized prices. In general, the existence of a martingale measure for “normalized” price processes is a necessary and sufficient condition for arbitrage opportunities not to be present among the three primitive securities. On the other hand, the uniqueness of a martingale measure is a necessary and sufficient condition for the market to be *dynamically complete* in that any contingent security can be “replicated” by dynamically trading in the primitive securities.

This martingale connection of an arbitrage-free price system for primitive securities, made originally by Cox and Ross (1976) and Harrison and Kreps (1979), is the conerstone for the general theory of contingent claim valuation in continuous time, of which the option pricing theory of Black and Scholes (1973) is just an example.

5 Concluding Remarks

We have briefly overviewed three main subjects of modern financial economics. The intent is not to give complete details but rather to summarize major results. In four articles to appear in this

journal, we will give more in depth discussions of the three subjects and some practical applications of the ideas. More specifically, in the first article to appear, we will lay out in greater detail the mathematics of the portfolio frontier and the economics of the Capital Asset Pricing Models. The applications of the Capital Asset Pricing Model to capital budgeting and to money management will also be discussed.

In the second article to appear, instead of giving more details on the dynamic portfolio theory and the Intertemporal Capital Asset Pricing Models, we will develop more fully the theory of contingent claims valuation in continuous time. The Black and Scholes (1973) option pricing theory will be the application to be discussed.

In the third article to appear, we will show how the theory of contingent claim valuation gives more insights into the dynamic portfolio theory and the intertemporal asset pricing models. Finally, in the fourth article to appear we will combine the valuation theory of contingent securities and the intertemporal asset pricing models to develop a model of the term structure of interest rates and apply the model to study the pricing of interest rate contingent securities.

6 References

1. Black, F. 1972. Capital market equilibrium with restricted borrowing. *Journal of Business* 45:444-454.
2. Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81:637-654.
3. Breeden, D. 1979. An intertemporal capital pricing model with stochastic investment opportunities. *Journal of Financial Economics* 7: 265-296.
4. Cox, J., and C. Huang. 1986. A variational problem arising in financial economics. to appear in *Journal of Mathematical Economics*.
5. Cox, J., and C. Huang. 1989. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory* 49:33-83.
6. Cox, J., J. Ingersoll, and S. Ross. 1985. An intertemporal general equilibrium model of asset prices. *Econometrica* 53:363-384.
7. Cox, J., S. Ross. 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3:145-166.
8. Duffie, D., and C. Huang. 1985. Implementing Arrow-Debreu equilibria by continuous trading of few long-lived securities. *Econometrica* 53:1337-1356.
9. Duffie, D. 1986. Stochastic equilibria: Existence, spanning number, and the "no expected gain from trade" hypothesis. *Econometrica* 54:1161-1184.
10. Harrison, M., and D. Kreps. 1979. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* 20:381-408.

11. Harrison, M., and S. Pliska. 1981. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications* 11:215–260.
12. He, H., and N. Pearson. 1988b. Consumption portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. Mimeo. Sloan School of Management, Massachusetts Institute of Technology.
13. Huang, C. 1985a. Information structure and equilibrium asset prices. *Journal of Economic Theory* 35:33–71.
14. Huang, C. 1985b. Information structure and viable price systems. *Journal of Mathematical Economics* 13:215–240.
15. Huang, C. 1987. An intertemporal general equilibrium asset pricing model: The case of diffusion information. *Econometrica* 55:117–142.
16. Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47:13–37.
17. Markowitz, H. 1952. Portfolio Selection. *Journal of Finance* 7:77–91.
18. Merton, R. 1971. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory* 3:373–413.
19. Merton, R. 1973a. An intertemporal capital asset pricing model. *Econometrica* 41:867–887.
20. Merton, R. 1973b. Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4:141–183.
21. Modigliani, F., and M. Miller. 1958. The cost of capital, corporate finance and the theory of corporation finance. *American Economic Review* 48:261–297.
22. Mossin, J. 1966. Equilibrium in a capital asset market. *Econometrica* 35:768–783.
23. Pagès, H. 1989. Three Essays In Optimal Consumption. Unpublished Ph.D. thesis. Massachusetts Institutes of Technology.
24. Sharpe, W. 1964. Capital asset prices: A theory of capital market equilibrium under conditions of risk. *Journal of Finance* 19:425–442.
25. Tobin, J. 1958. Liquidity preferences as behavior towards risk. *Review of Financial Studies* 25:65–86.

The Static Portfolio Theory and the Capital Asset Pricing Model

Chi-fu Huang

1 Introduction

In the static portfolio theory, it is assumed that investors choose portfolios of assets according to the means and variances of the returns on the portfolios. In particular, investors prefer the mean and dislike the variance. This mean–variance model of asset choice has been used extensively in finance since its development by Markowitz (1952) almost three decades ago. The popularity of this model lies not in the generality but in the tractability that its assumptions afford. Moreover, the closing of this mean-variance model using the market clearing condition yields an equilibrium relation among the expected rates of return on risky assets, termed the *Capital Asset Pricing Model (CAPM)*, which, till these days, remains to be one of the most important asset pricing models.

The purpose of this paper is just to give a brief account of the static portfolio theory and its implications on the equilibrium expected asset return. We will also give two applications of the CAPM, one in corporate finance and one in investments. The rest of this paper is organized as follows. Section 2 lays out the basic static portfolio theory. Section 3 presents some deeper results in portfolio theory in preparation for the introduction of CAPM. Section 4 derives the CAPM. Applications of the CAPM are in Section 5. Section 6 contains the concluding remarks.

2 Static Portfolio Theory

The static portfolio theory characterizes a set of portfolios from which investors choose their optimal portfolios. It is assumed that investors prefer the mean and dislike the variance of the rate of return on a portfolio. Thus the set of portfolios from which optimal portfolios will be chosen is composed of those portfolios that have the smallest variance of the rate of return for given levels of the expected rate of return. We will show that this set, termed *portfolio frontier*, is a parabola in the variance-expected rate of return space and is a hyperbola in the standard deviation-expected rate of return space. One of the fundamental results in the portfolio theory is that the portfolio frontier has the *two-fund separation* property: a frontier portfolio is a portfolio of any two distinct frontier portfolios. This property has a significant implication in the market for mutual funds: we only need two mutual funds that are frontier portfolios since all the portfolios that are possible optimal choices for investors are portfolios of these two frontier portfolios.

We now proceed to derive this two-fund separation property. We suppose that there are $N \geq 2$ risky assets traded in a frictionless economy. The rates of return on these assets have finite variances and unequal expectations. Assumed also that asset returns are linearly independent and thus their variance-covariance matrix $\mathbf{V}$ is nonsingular. Note that by the fact that $\mathbf{V}$ is a variance-covariance matrix, it is symmetric and positive definite. The rate of return on i -th asset is denoted $\tilde{r}_i$ and the N -vector of asset rates of return is denoted by $\tilde{\mathbf{r}}$.

A portfolio is said to be a *frontier portfolio* if it has the minimum variance among portfolios that have the same expected rate of return. A portfolio p is a frontier portfolio if and only if $\mathbf{w}_p$, the N -vector portfolio weights of p , is the solution to the quadratic program:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \frac{1}{2} \mathbf{w}^\top \mathbf{V} \mathbf{w} \\ \text{s.t.} \quad & \mathbf{w}^\top \mathbf{e} = E[\tilde{r}_p], \\ & \mathbf{w}^\top \mathbf{1} = 1, \end{aligned} \tag{16}$$

where $\mathbf{e}$ denotes the N -vector of expected rates of return on the N risky assets and $E[\tilde{r}_p]$ denotes the expected rate of return on portfolio p .

Forming the Lagrangian, $\mathbf{w}_p$ is the solution to the following:

$$\min_{\mathbf{w}, \lambda, \gamma} L = \frac{1}{2} \mathbf{w}^\top \mathbf{V} \mathbf{w} + \lambda (E[\tilde{r}_p] - \mathbf{w}^\top \mathbf{e}) + \gamma (1 - \mathbf{w}^\top \mathbf{1}),$$

where λ and γ are two positive constants. The first order conditions are,

$$\frac{\partial L}{\partial \mathbf{w}} = \mathbf{V} \mathbf{w}_p - \lambda \mathbf{e} - \gamma \mathbf{1} = \mathbf{0}, \tag{17}$$

$$\frac{\partial L}{\partial \lambda} = E[\tilde{r}_p] - \mathbf{w}_p^\top \mathbf{e} = 0, \tag{18}$$

$$\frac{\partial L}{\partial \gamma} = 1 - \mathbf{w}_p^\top \mathbf{1} = 0, \tag{19}$$

where $\mathbf{1}$ denotes an $N \times 1$ vector of one's. Since $\mathbf{V}$ is a positive definite matrix, it follows that the first order conditions are necessary and sufficient for a global optimum.

Solving (17) for $\mathbf{w}_p$ gives,

$$\mathbf{w}_p = \lambda (\mathbf{V}^{-1} \mathbf{e}) + \gamma (\mathbf{V}^{-1} \mathbf{1}). \tag{20}$$

Premultiplying both sides of relation (20) by $\mathbf{e}^\top$ and using (18) gives,

$$E[\tilde{r}_p] = \lambda (\mathbf{e}^\top \mathbf{V}^{-1} \mathbf{e}) + \gamma (\mathbf{e}^\top \mathbf{V}^{-1} \mathbf{1}). \tag{21}$$

Premultiplying both sides of relation (20) by $\mathbf{1}^\top$ and using (19) gives,

$$1 = \lambda (\mathbf{1}^\top \mathbf{V}^{-1} \mathbf{e}) + \gamma (\mathbf{1}^\top \mathbf{V}^{-1} \mathbf{1}). \tag{22}$$

Solving (21) and (22) for λ and γ gives,

$$\begin{aligned} \lambda &= \frac{CE[\tilde{r}_p] - A}{D} \\ \gamma &= \frac{B - AE[\tilde{r}_p]}{D} \end{aligned}$$

where

$$\begin{aligned} A &= \mathbf{1}^\top \mathbf{V}^{-1} \mathbf{e} = \mathbf{e}^\top \mathbf{V}^{-1} \mathbf{1}, \\ B &= \mathbf{e}^\top \mathbf{V}^{-1} \mathbf{e}, \\ C &= \mathbf{1}^\top \mathbf{V}^{-1} \mathbf{1}, \\ D &= BC - A^2. \end{aligned}$$

Since $\mathbf{V}^{-1}$ is positive definite, $B > 0$ and $C > 0$. We claim that $D > 0$. To see this, we note that,

$$(\mathbf{Ae} - B\mathbf{1})^\top \mathbf{V}^{-1} (\mathbf{Ae} - B\mathbf{1}) = B(BC - A^2).$$

Since the left-handside of the above relation is strictly positive, we have $BC - A^2 > 0$, or $D > 0$.

Substituting for λ and γ in relation (20) gives the unique set of portfolio weights for the frontier portfolio having an expected rate of $E[\tilde{r}_p]$:

$$\mathbf{w}_p = \mathbf{g} + hE[\tilde{r}_p], \quad (23)$$

where

$$\mathbf{g} = \frac{1}{D}[B(\mathbf{V}^{-1}\mathbf{1}) - A(\mathbf{V}^{-1}\mathbf{e})]$$

and

$$h = \frac{1}{D}[C(\mathbf{V}^{-1}\mathbf{e}) - A(\mathbf{V}^{-1}\mathbf{1})].$$

We can verify quickly that $\mathbf{g}$ is the vector of portfolio weights corresponding to a frontier portfolio having a zero expected rate of return and that $\mathbf{g} + h$ is the vector of portfolio weights of a frontier portfolio having an expected rate of return equal to 1. Note that (23) is a necessary and sufficient condition for $\mathbf{w}_p$ to be the vector of portfolio weights for a frontier portfolio.

Relation (23) immediately implies the two-fund separation property. To see this, let p_1 and p_2 be two distinct frontier portfolios and let q be any frontier portfolio. Since $E[\tilde{r}_{p_1}]$ is not equal to $E[\tilde{r}_{p_2}]$, there exists a unique real number α such that,

$$E[\tilde{r}_q] = \alpha E[\tilde{r}_{p_1}] + (1 - \alpha)E[\tilde{r}_{p_2}].$$

Now consider a portfolio of p_1 and p_2 with weights $\{\alpha, (1 - \alpha)\}$. We have

$$\begin{aligned} \alpha \mathbf{w}_{p_1} + (1 - \alpha) \mathbf{w}_{p_2} &= \alpha(\mathbf{g} + hE[\tilde{r}_{p_1}]) + (1 - \alpha)(\mathbf{g} + hE[\tilde{r}_{p_2}]) \\ &= \mathbf{g} + hE[\tilde{r}_q] \\ &= \mathbf{w}_q. \end{aligned}$$

Thus the entire portfolio frontier can be *generated* by any two distinct frontier portfolios.

Conversely, (23) also implies that any portfolio of two frontier portfolios is still a frontier portfolio; that is, the portfolio frontier is a linear space. To see this, let α be any given scalar. Then the analysis in the previous paragraph shows that $\alpha \mathbf{w}_{p_1} + (1 - \alpha) \mathbf{w}_{p_2}$ can be written in the form of (23) and thus is a vector of portfolio weights that generate a frontier portfolio. A straightforward generalization of this argument shows that a portfolio of a finite number of frontier portfolios is also a frontier portfolio.

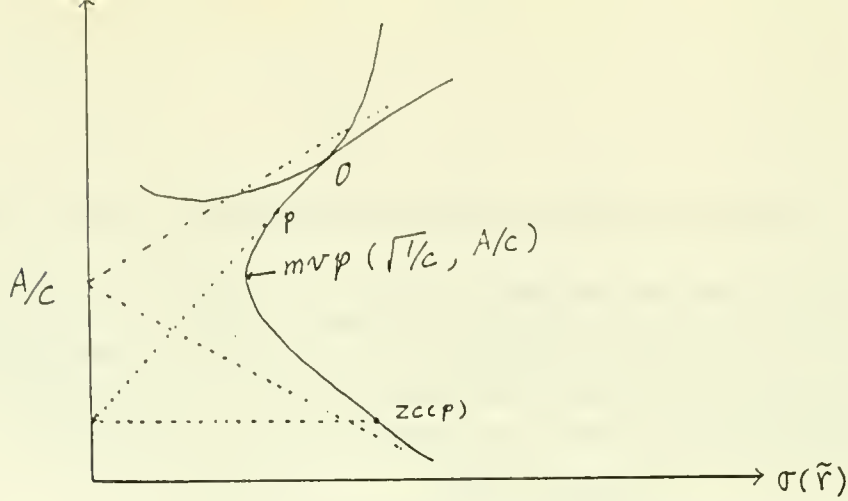


Figure 3: A Portfolio Frontier without a Riskless Asset

Now that we have characterized the set of portfolios from which optimal portfolios will be chosen, we turn our attention to how a given investor makes his or her portfolio choices. In preparation for a geometric presentation of the ideas, we first analyze in some detail the geometric properties of the portfolio frontier.

By definition, the covariance between the rates of return on any two frontier portfolios p and q is,

$$\text{Cov}(\tilde{r}_p, \tilde{r}_q) = \mathbf{w}_p^T \mathbf{V} \mathbf{w}_q = \frac{C}{D} (E[\tilde{r}_p] - A/C)(E[\tilde{r}_q] - A/C) + \frac{1}{C}. \quad (24)$$

Using the definition of the variance of the rate of return of a portfolio and (24) gives

$$\frac{\sigma^2(\tilde{r}_p)}{1/C} - \frac{(E[\tilde{r}_p] - A/C)^2}{D/C^2} = 1, \quad (25)$$

which is a parabola in the variance-expected rate of return space and is a hyperbola in the standard deviation-expected rate of return space with center $(0, A/C)$ and asymptotes $E[\tilde{r}_p] = A/C \pm \sqrt{D/C} \sigma(\tilde{r}_p)$, where $\sigma(\tilde{r}_p)$ denotes the standard deviation of the return on the portfolio p . Figure 3 depicts this hyperbola in the standard deviation-expected rate of return space.

The portfolio having the minimum variance of all possible portfolios, or the *minimum variance portfolio (mvp)*, is at $(\sqrt{1/C}, A/C)$ in Figure 3. The reader may have noticed by now that since investors prefer expected rate of return and dislike variance, they will never choose any frontier portfolio that has an expected rate of return less than A/C . This follows since for any such portfolio, there exists another frontier portfolio that has the same variance and has a strictly higher expected rate of return. Thus an investor will never choose the former over the latter. We will therefore call those frontier portfolios above *mvp* the *efficient frontier portfolios*. In contrast, those frontier portfolios below *mvp* will be called *inefficient frontier portfolios*. Using (23), it is easily shown that any convex combination of (in)efficient frontier portfolios remains to be an (in)efficient frontier portfolio. Thus the set of (in)efficient portfolios is a convex set.

Since an investor's preferences can be captured by his or her preferences for the first two moments and he or she prefers expected rate of return and dislikes variance, his or her preferences can be represented by upward sloping indifference curves in Figure 3. Note that an indifference curve traces through the combinations of standard deviations and expected rates of return that give the investor the same level of satisfaction. The indifference curve that is most further to the northwest gives the highest level of satisfaction. The optimal portfolio for an investor is then determined by

the tangent point between his indifference curves and the portfolio frontier as indicated by o in Figure 3.

The above analysis assumes that there does not exist a riskless asset. When there exists a riskless asset, we can simplify our analysis a bit. Still let $\mathbf{w}$ denote the N -vector portfolio weights on risky assets. Then $(1 - \mathbf{w}^\top \mathbf{1})$ is the portfolio weight on the riskless asset. Define the portfolio frontier as in the case of all risky assets. Then $\mathbf{w}_p$ is an N -vector of portfolio weights on risky assets for a frontier portfolio having expected rate of return $E[\tilde{r}_p]$ if it is a solution to

$$\begin{aligned} \min_{\mathbf{w}} \quad & \frac{1}{2} \mathbf{w}^\top \mathbf{V} \mathbf{w} \\ \text{s.t.} \quad & \mathbf{w}^\top \mathbf{e} + (1 - \mathbf{w}^\top \mathbf{1}) r_f = E[\tilde{r}_p], \end{aligned} \quad (26)$$

where r_f is the rate of return on the riskless asset. Forming the Lagrangian, the first order necessary and sufficient conditions for $\mathbf{w}_p$ to be the solution to (26) are

$$\mathbf{V} \mathbf{w}_p = \lambda (\mathbf{e} - r_f \mathbf{1})$$

and

$$r_f + \mathbf{w}_p^\top (\mathbf{e} - r_f \mathbf{1}) = E[\tilde{r}_p].$$

Solving for $\mathbf{w}_p$, we have

$$\mathbf{w}_p = \mathbf{V}^{-1} (\mathbf{e} - r_f \mathbf{1}) \frac{E[\tilde{r}_p] - r_f}{H}, \quad (27)$$

where $H = (\mathbf{e} - r_f \mathbf{1})^\top \mathbf{V}^{-1} (\mathbf{e} - r_f \mathbf{1}) = B - 2r_f A + r_f C$. The variance of the rate of return on portfolio p is,

$$\sigma^2(\tilde{r}_p) = \mathbf{w}_p^\top \mathbf{V} \mathbf{w}_p = \frac{(E[\tilde{r}_p] - r_f)^2}{H}. \quad (28)$$

Thus, the portfolio frontier of all asset is composed of two half-lines emanating from the point $(0, r_f)$ with slope $\sqrt{H}$ and $-\sqrt{H}$. The half-line with a positive slope is the efficient portfolio frontier; while the half-line with a negative slope is the inefficient portfolio frontier. One can show that if $r_f < A/C$, which is the case of interest, the efficient portfolio frontier is the half-line emanating from $(0, r_f)$ that is tangent to the portfolio frontier of risky assets; see Figure 4. The frontier portfolios above the tangent portfolio e are constructed by borrowing at the riskless asset and investing this fund together with the initial wealth in the tangent portfolio e . The frontier portfolios between mvp and e are convex combinations of them. And, the frontier portfolios below mvp are constructed by short-selling the tangent portfolio e and investing the proceeds from the short-sale and the initial wealth in the riskless asset.

The optimal portfolio for an investor is again determined by the tangent point between an indifference curve and the portfolio frontier; see Figure 4.

This concludes our discussion on the basics of the static portfolio theory.

3 Further Results on the Mathematics of the Portfolio Frontier

We report in this section some more mathematical properties of the portfolio frontier in preparation for introducing the *Capital Asset Pricing Model (CAPM)*. A clear understanding of these properties will help us in the next section focus on the economics of the CAPM.

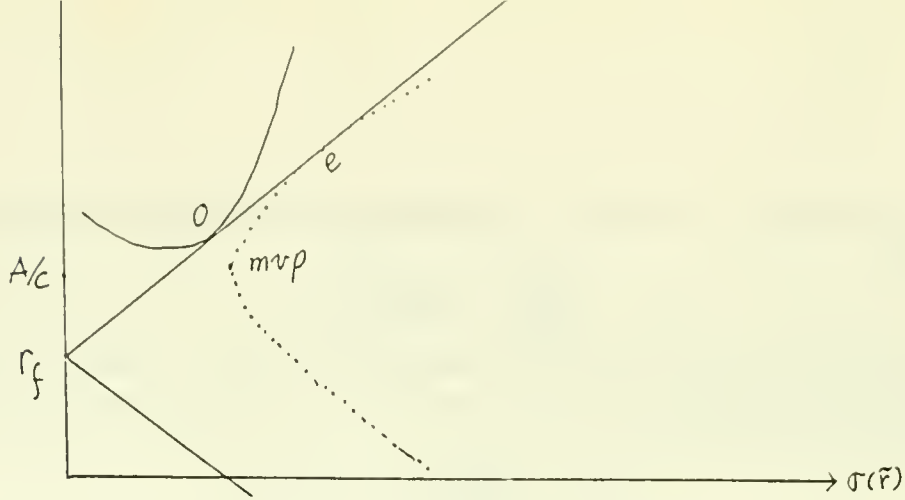


Figure 4: A Portfolio Frontier with a Riskless Asset

First, consider the case where there is no riskless asset. We start by recording a special property of the *mvp*: the covariance of the rate of return on *mvp* and that on any portfolio (not necessarily on the frontier) is always equal to the variance of the rate of return on *mvp*. To see this, consider a portfolio of *mvp* and any portfolio *p* with weights $1 - a$ and a and with minimum variance. Since *mvp* has the minimum variance among all portfolios, $a = 0$ must be the solution to the following program:

$$\min_a a^2 \sigma^2(\tilde{r}_p) + 2a(1 - a)\text{Cov}(\tilde{r}_p, \tilde{r}_{mvp}) + (1 - a)^2 \sigma^2(\tilde{r}_{mvp}).$$

The first order necessary and sufficient condition for $a = 0$ to be the solution is:

$$\text{Cov}(\tilde{r}_p, \tilde{r}_{mvp}) = \sigma^2(\tilde{r}_{mvp}) = \frac{1}{C}, \quad (29)$$

which was to be proved.

Next we introduce two key properties that will be useful to our derivation of the CAPM. First, for any portfolio *p* on the frontier, except for *mvp*, there exists a unique frontier portfolio, denoted by *zc(p)*, which has a zero covariance with *p*. Second, the expected rate of return on any portfolio has an exact linear relation with the expected rates of return on a frontier portfolio $p \neq mvp$ and on *zc(p)*.

Setting the covariance between two frontier portfolios *p* and *zc(p)*, given in relation (24), equal to zero. Solving for the expected rate of return on *zc(p)*, we get :

$$E[\tilde{r}_{zc(p)}] = A/C - \frac{D/C^2}{E[\tilde{r}_p] - A/C}. \quad (30)$$

It is confirmed from relation (30) that there does not exist a frontier portfolio that has zero covariance with *mvp* as evidenced in (29).

Equation (30) gives us a clue to the location of *zc(p)*. If *p* is an efficient frontier portfolio, then (30) implies that

$$E[\tilde{r}_{zc(p)}] < A/C.$$

Thus *zc(p)* is an inefficient frontier portfolio. Similarly, if *p* is an inefficient frontier portfolio, the *zc(p)* is an efficient frontier portfolio. Geometrically *zc(p)* can be located by the following fact: in the standard deviation-expected rate of return space, the intercept on the expected rate of return axis of the line tangent to the portfolio frontier at the point associated with any frontier portfolio

p is $E[\tilde{r}_{zc(p)}]$; see Figure 3. To see this, we first differentiate (25) totally with respect to $\sigma(\tilde{r}_p)$ and $E[\tilde{r}_p]$ to obtain

$$\frac{dE[\tilde{r}_p]}{d\sigma(\tilde{r}_p)} = \frac{\sigma(\tilde{r}_p)D}{CE[\tilde{r}_p] - A},$$

which is the slope of the portfolio frontier at the point $(\sigma(\tilde{r}_p), E[\tilde{r}_p])$. The expected rate of return axis intercept of the tangent line is just

$$E[\tilde{r}_p] - \frac{dE[\tilde{r}_p]}{d\sigma(\tilde{r}_p)}\sigma(\tilde{r}_p) = E[\tilde{r}_{zc(p)}].$$

Now we consider the relationship between the expected rate of return on any portfolio q (not necessarily on the frontier), and those on the frontier portfolios. Let p be a frontier portfolio other than the mvp , and let q be any portfolio. The covariance of $\tilde{r}_p$ and $\tilde{r}_q$ is

$$\begin{aligned} \text{Cov}(\tilde{r}_p, \tilde{r}_q) &= \mathbf{w}_p^\top \mathbf{V} \mathbf{w}_q \\ &= \lambda E[\tilde{r}_q] + \gamma. \end{aligned} \quad (31)$$

Substituting λ and γ into (31) gives,

$$E[\tilde{r}_q] = E[\tilde{r}_{zc(p)}] + \beta_{qp}(E[\tilde{r}_p] - E[\tilde{r}_{zc(p)}]), \quad (32)$$

where $\beta_{qp} \equiv \text{Cov}(\tilde{r}_q, \tilde{r}_p) / \sigma^2(\tilde{r}_p)$. Equation (32) says that, for a fixed frontier portfolio p which is not mvp , there exists a linear relation between the expected rate of return of any portfolio q and its “beta” with respect to any frontier portfolio $p \neq mvp$, β_{qp} . If p is on the efficient frontier, then the higher the β_{qp} , the larger the expected rate of return on q in excess of that on p , since $E[\tilde{r}_p] > E[\tilde{r}_{zc(p)}]$.

Next we consider the case where there exists a riskless asset with a rate of return r_f . A relation similar to (32) can be derived. Let q be any portfolio, with $\mathbf{w}_q$ the portfolio weights on the N risky assets. Also, let $\mathbf{w}_p$ be the portfolio weights on the N risky assets of a frontier portfolio p . We assume that $E[\tilde{r}_p] \neq r_f$. Then,

$$\begin{aligned} \text{Cov}(\tilde{r}_q, \tilde{r}_p) &= \mathbf{w}_q^\top \mathbf{V} \mathbf{w}_p \\ &= \frac{(E[\tilde{r}_q] - r_f)(E[\tilde{r}_p] - r_f)}{H}. \end{aligned}$$

Using (28), we obtain

$$E[\tilde{r}_q] = r_f + \beta_{qp}(E[\tilde{r}_p] - r_f). \quad (33)$$

That is, there exists a linear relation between the expected rate of return on any portfolio q and its “beta” with respect to a frontier portfolio $p \neq mvp$.

Note that (32) and (33) are mathematical relations. Given a collection of random variable with finite variances, these two relations will always be valid.

4 The Capital Asset Pricing Model

Now we are ready to present the implication of the static portfolio theory on the equilibrium expected rates of return on assets, or the CAPM. Assume throughout that investors prefer to hold efficient frontier portfolios and that if there exists a riskless asset, it is in zero net supply.

Let $W_0^i > 0$ be investor i 's initial wealth, and let w_{ij} be the optimal proportion of the initial wealth invested in the j -th risky asset by investor i . The total wealth in the economy is

$$W_{m0} \equiv \sum_{i=1}^I W_0^i,$$

where I is the total number of investors in the economy. In the market equilibrium, the total wealth W_{m0} is equal to the total value of assets. Since the riskless asset, if it exists, is in zero net supply, the total value of assets is equal to the total value of risky assets. Let w_{mj} denote the proportion of the total wealth contributed by the total value of j -th risky security. For markets to clear, we must then have

$$\sum_{i=1}^I w_{ij} W_0^i = w_{mj} W_{m0}.$$

This relation is equivalent to

$$\sum_{i=1}^I w_{ij} \frac{W_0^i}{W_{m0}} = w_{mj},$$

where we note that $\frac{W_0^i}{W_{m0}} > 0$ and

$$\sum_{i=1}^I \frac{W_0^i}{W_{m0}} = 1.$$

Thus, the portfolio weights for the *market portfolio*, w_{mj} 's, are a convex combination of the optimal portfolio weights for investors in an market equilibrium.

We take two cases. Case 1: there is no riskless asset. By the hypothesis that investors prefer to hold efficient frontier portfolios, the market portfolio, being a convex combination of investors' optimal portfolios in equilibrium, is an efficient frontier portfolio. By the two-fund separation property of the portfolio frontier, the two separating mutual funds in equilibrium can be chosen, for example, to be the market portfolio and its zero-covariance portfolio.

Moreover, since the market portfolio is on the portfolio frontier, (32) implies that, for any portfolio q ,

$$E[\tilde{r}_q] = E[\tilde{r}_{zc(m)}] + \beta_{qm}(E[\tilde{r}_m] - E[\tilde{r}_{zc(m)}]), \quad (34)$$

where

$$\tilde{r}_m = \sum_{j=1}^N w_{mj} \tilde{r}_j$$

is the rate of return on the market portfolio, and

$$\beta_{qm} = \frac{\text{Cov}(\tilde{r}_q, \tilde{r}_m)}{\text{Var}(\tilde{r}_m)}.$$

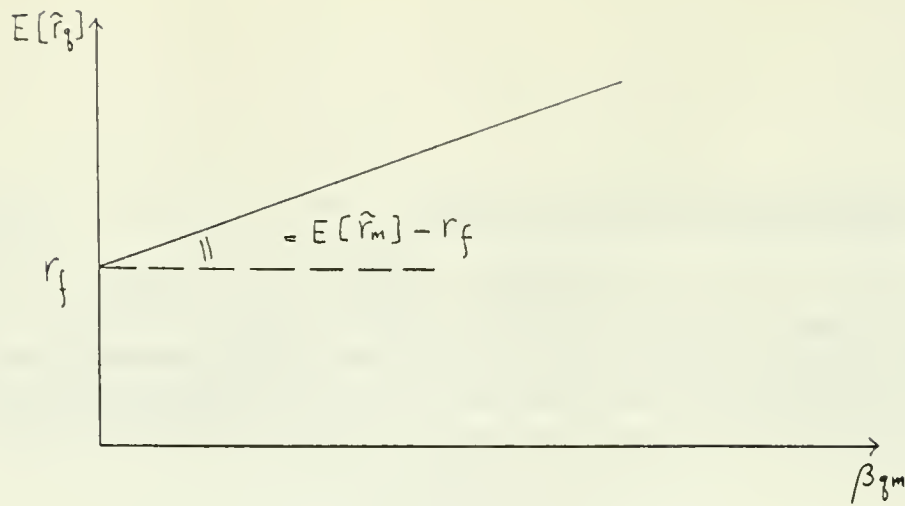


Figure 5: The Security Market Line

Note that since the market portfolio is an efficient frontier portfolio, $E[\tilde{r}_m] - E[\tilde{r}_{zc(m)}] > 0$. Thus, the higher the “beta” of q with respect to the market portfolio, the larger the expected rate of return on q in excess of that on the market portfolio. Relation (34) is known to be the *Zero-Beta Capital Asset Pricing Model*, due to Black (1972) and Lintner (1969).

Case 2: there is a riskless asset. First, the reader should convince himself that if $r_f > A/C$, then the inefficient portfolio frontier is the half-line emanating from r_f that is tangent to the portfolio frontier of only risky assets. On the other hand, the efficient portfolio frontier is the half-line from r_f that has a slope $\sqrt{H}$. Thus all the efficient frontier portfolios are constructed by short-selling the tangent portfolio between the inefficient portfolio frontier of all assets and that of risky assets only and investing the proceeds of the short-sale and the initial wealth in the riskless asset. Since the riskless asset is in zero net supply, this scenario cannot be sustained in a market equilibrium as every investor will demand a positive amount of the risk asset. Similar arguments can be made when $r_f = A/C$. Thus the only situation consistent with a market equilibrium is $r_f < A/C$ as depicted in Figure 4. From Figure 4 we know that investors will all be optimally holding a portfolio of the tangent portfolio e and the riskless asset. The tangent portfolio, being on the portfolio frontier of risky assets, is composed only of risky assets. Since e is the single risky asset portfolio that every investor holds in equilibrium, for the markets to clear, it is necessary that e is the market portfolio. Thus in equilibrium, all investors hold portfolios of the market portfolio and the riskless asset. The efficient portfolio frontier in equilibrium is termed the *capital market line*.

In addition, (33) immediately implies that

$$E[\tilde{r}_q] = r_f + \beta_{qm}(E[\tilde{r}_m] - r_f). \quad (35)$$

This is the *Capital Asset Pricing Model* (CAPM) independently derived by Lintner (1965), Mossin (1965), and Sharp (1964). The plot of (35) in the beta-expected rate of return space is termed the *security market line*; see Figure 5.

To simplify our discussion, we will assume henceforth that there exists a riskless asset.

In a world of uncertainty, investors will be rewarded for the risk they bear by investing in risky assets. In the market equilibrium, the reward and risk relation must be determined so that all risky assets are equally attractive or else investors will find an reallocation of their investments in the risky assets to be desirable and this is inconsistent with an equilibrium. In order for all assets to be equally attractive, riskier assets should therefore yield higher rewards.

The reward for asset q in the context of the CAPM is $E[\tilde{r}_q] - r_f$, interpreted to be the “risk premium” on asset q . The measure of risk, however, is the beta of an asset with respect to the

market portfolio. Also, the reward-risk relation is linear as described in (35). Thus the higher the “market beta”, β_{qm} , of portfolio q , the higher the risk premium on portfolio q in equilibrium. In other words, the higher the market beta of an asset, the riskier the asset is as viewed by investors and thus they demand a higher risk premium.

The randomness of an asset’s return that is correlated with that of the market portfolio is termed its *systematic risk* and the residual is termed the *unsystematic risk*. A message of the CAPM is that an investors will not be rewarded for bearing unsystematic risks. This is so, because an investor can diversify away the unsystematic risks by forming a large portfolio. In fact, in equilibrium, all investors hold a portfolio of the market portfolio and the riskless asset and no investor is bearing any unsystematic risk. This is the essential reason that an asset’s market beta is the proper measure of its risk.

Risky assets whose returns are positively correlated with the return on the market portfolio have positive risk premiums. In such event, the higher the asset beta, the higher the risk premium. On the other hand, assets whose returns are negatively correlated with the return on the market portfolio will have negative risk premium. The intuition of this relationship can be understood as follows. Consider two assets A and B . Asset A and asset B have the same expected time-1 payoffs. However, asset A ’s payoff are positively correlated with the payoffs of the market portfolio, while asset B ’s payoffs are negatively correlated with that of the market portfolio. That is, asset A has high payoffs when the economy is in relatively prosperous states, while asset B has high payoffs when the economy is relatively poor states. One unit of the payoff is more valuable in a relatively poor state than in a relative abundant state. Therefore, asset B is more desirable, and its price will be higher than that of asset A . Since assets A and B have the same expected payoffs, the expected rate of return of asset A will be higher than that for asset B . In other words, asset A ’s payoff structure is not as attractive as that of asset B . Therefore, it has to yield a higher expected rate of return than asset B to make itself as attractive as asset B in equilibrium.

Before leaving this section, we caution the reader to note that the economics of the CAPM lies not in the linear relation among expected asset returns but in the identification of the market portfolio to be on the efficient portfolio frontier. From the mathematics of the portfolio frontier, the linear relation (32) holds for all portfolio q once a frontier portfolio p is identified. In the CAPM, the market portfolio is identified to be on the efficient portfolio frontier using the market clearing condition. This is the most important economics of the model.

5 The Application of the Capital Asset Pricing Model

We present two applications of the CAPM in this section. The first is in corporate finance regarding projects selection and the second is in investments regarding securities selection.

5.1 Capital budgeting

Let the random payoff of a risky asset be $\tilde{y}$, and let S_y be its current equilibrium price. Suppose that the CAPM holds. By definition, we know $\bar{r}_y = \tilde{y}/S_y - 1$. Using this definition and (33) we

can write

$$S_y = \frac{E[\tilde{y}] - \phi^* \rho_{ym} \sigma(\tilde{y})}{1 + r_f}, \quad (36)$$

where

$$\phi^* \equiv \frac{E[\tilde{r}_m] - r_f}{\sigma(\tilde{r}_m)}$$

is the slope of the capital market line and

$$\rho_{ym} = \frac{\text{Cov}(\tilde{y}, \tilde{r}_m)}{\sigma(\tilde{y})\sigma(\tilde{r}_m)}.$$

This relation gives us a way to calculate the “fair price” or “present value” of a risky asset with a payoff represented by $\tilde{y}$. The numerator can be viewed as the “certainty equivalent” of the uncertain payoff $\tilde{y}$. The present value of $\tilde{y}$ is equal to its certainty equivalent discounted at one plus the riskless rate.

Now observe the following. Let $\tilde{x}$ be the random payoff of another asset and let S_x be its equilibrium value. Consider the combination of $\tilde{x}$ and $\tilde{y}$ and let $\tilde{z}$ and S_z denote this combination and its current value, respectively. We will say that there is no synergy of this combination if $\tilde{z} = \tilde{x} + \tilde{y}$. On the other hand, this combination creates synergy if $\tilde{z} > \tilde{x} + \tilde{y}$. When there is no synergy, it is easily verified that

$$\begin{aligned} S_z &= \frac{E[\tilde{x} + \tilde{y}] - \phi^* \rho_{(x+y)m} \sigma(\tilde{x} + \tilde{y})}{1 + r_f} \\ &= \frac{E[\tilde{x}] - \phi^* \rho_{xm} \sigma(\tilde{x})}{1 + r_f} + \frac{E[\tilde{y}] - \phi^* \rho_{ym} \sigma(\tilde{y})}{1 + r_f} \\ &= S_x + S_y. \end{aligned}$$

That is, value of $\tilde{x} + \tilde{y}$ is equal to the sum of values of $\tilde{x}$ and $\tilde{y}$, or values are additive.

The additivity of values in the absence of synergy in a frictionless economy has one important implication for capital budgeting. Since the proper objective of a firm in a frictionless economy is to maximize its value and since values are additive, a firm should select all the projects whose present values of future payoffs exceed the initial costs. For example, let a project be represented by its initial cost I_0 and a future random payoff $\tilde{x}$. A firm facing this project should accept this project if and only if $S_x \geq I_0$ or the “net present value” of this project is positive. Note that this criterion for project selection is independent of the existing assets of the firm. A firm should not accept a project that has a negative net present value for the purpose of diversification. A project should be evaluated on its own merit using the net present value criterion.

Applying the same principle of value additivity to mergers and acquisitions leads to the conclusion that these activities do not create value unless there exists synergy – that is, unless the payoffs of the merged firms are greater than the sum of the payoffs of the individual firms.

The CAPM is essentially a two period model. Thus our discussion above on capital budgeting is limited to capital projects that last for two periods: one has to make a decision at time $t = 0$ about whether to undertake a project by incurring a capital outlay now for I_0 in exchange for a random cash flow of $\tilde{x}$ at time $t = 1$. In reality, however, capital investment projects last sometimes for up to twenty to thirty years. The conditions under which the net present value analysis using

CAPM can be extended to multiperiod economies are outside the scope of our discussion here, to which we refer interested readers to Myers and Turnbull (1977), for example.

5.2 Security Selection

The analysis of investments in securities is similar to that in capital budgeting. If all securities are priced fairly, then the optimal portfolio for an investor is a mixture of the market portfolio and the riskless asset. However, if there are mispriced securities in that the present value of a security is different from its current market price, then one should take advantage of this mispricing. One can identify a mispriced security by conducting the present value analysis described in section 5.1. Here, however, we will present an alternative method which is especially suited for investments in securities.

Let I_0 be the current market price of a security paying $\tilde{x}$ next period and let S_x be its present value determined according to (36). Let the rate of return of investing in x be denoted by

$$\tilde{r}_x^p = \frac{\tilde{x}}{I_0} - 1.$$

By the CAPM, we know

$$E\left[\frac{\tilde{x}}{S_x}\right] = 1 + r_f + \beta_{xm}(E[\tilde{r}_m] - r_f),$$

where we recall that

$$\beta_{xm} \equiv \frac{\text{Cov}(\tilde{x}/S_x, \tilde{r}_m)}{\sigma^2(\tilde{r}_m)}.$$

Substituting the definition of $\tilde{r}_x^p$ into the above relation gives

$$E[\tilde{r}_x^p] - r_f = \alpha_x + \beta_{xm}^p(E[\tilde{r}_m] - r_f), \quad (37)$$

where

$$\alpha_x \equiv \left(\frac{S_x}{I_0} - 1\right)(1 + r_f)$$

and

$$\beta_{xm}^p \equiv \frac{\text{Cov}(\tilde{x}/I_0, \tilde{r}_m)}{\sigma^2(\tilde{r}_m)}.$$

Note that the asset x is fairly priced if $I_0 = S_x$, it is over-priced if and only if $I_0 > S_x$, and it is under-priced if and only if $I_0 < S_x$. These and (37) imply that the asset is

fairly priced	if and only if	$\alpha_x = 0$;
over-priced	if and only if	$\alpha_x < 0$;
under-priced	if and only if	$\alpha_x > 0$.

Now we turn our attention to how we empirically identify mispriced securities. First, we assume that the CAPM holds period by period in a multiperiod economy. Second, assume that return distributions are stationary. Let $\tilde{r}_{jt}$ denote the rate of return from time $t - 1$ to t from investing in security x . Similarly let $\tilde{r}_{mt}$ denote the rate of return from $t - 1$ to t on the market portfolio. Suppose we are at time t and have observed realized rates of return on security x : $\tilde{r}_{x1}, \tilde{r}_{x2}, \dots, \tilde{r}_{xt}$,

and on the market portfolio $\bar{r}_{m1}, \bar{r}_{m2}, \dots, \bar{r}_{mt}$. By the hypothesis that return distributions are stationary, we can estimate α_x and β_{xm}^p by running the regression of $\bar{r}_{xs} - r_{fs}$ on $\bar{r}_{ms} - r_{fs}$:

$$\bar{r}_{xs} - r_{fs} = \hat{\alpha}_x + \hat{\beta}_{xm}^p(\bar{r}_{ms} - r_{fs}) + \tilde{\epsilon}_{xs}, \quad s = 1, 2, \dots, t,$$

where r_{fs} denotes the riskless interest rate from time $s-1$ to s and $\hat{\alpha}_x$ and $\hat{\beta}_{xm}^p$ denote the estimates of α_x and β_{xm}^p , respectively. If $\hat{\alpha}_x$ is significantly different from zero, then we say that security x is not fairly priced according to the CAPM. Otherwise, x is fairly priced. When $\hat{\alpha}_x$ is significantly greater (smaller) than zero, security x is under-priced (over-priced). Note that $\tilde{\epsilon}_{xs}$ is the risk at time s that is uncorrelated with the return on the market portfolio or the unsystematic risk. To simplify notation, we will use $\hat{\cdot}$ to denote an estimate of a quantity. For example, the estimate of the standard deviation of the unsystematic risk will be denoted by $\hat{\sigma}(\tilde{\epsilon}_x)$. One exception is that we will use $\bar{r}_x$ and $\bar{r}_m$ to denote the estimates of the expected rates of return on x and on m , respectively.

The next question is: after a mispriced security is identified, say an under-priced security, should an investor invest all his money in this security? The answer is clearly no as investing solely in the security will expose the investor to too much unsystematic risk. On the other hand, as we will show now, one can construct a portfolio of this mispriced security, the market portfolio, and the riskless asset that has the same expected rate of return as the market portfolio but has a strictly lower variance.

Let w_x and w_m be the portfolio weights on x and on the market portfolio, respectively. Then the expected rate of return on this portfolio is $w_x \bar{r}_x + w_m \bar{r}_m + (1 - w_x - w_m)r_{f,t+1}$. Setting this to be equal to $\bar{r}_m$ gives

$$w_m = 1 - \frac{\hat{\alpha}_x + \hat{\beta}_{xm}^p(\bar{r}_m - r_{f(t+1)})}{\bar{r}_m - r_{f(t+1)}} w_x. \quad (38)$$

This gives the relation between w_x and w_m for a portfolio that has an expected rate of return equal to $\bar{r}_m$. The variance of this portfolio is

$$\left(1 - \frac{\hat{\alpha}_x}{\bar{r}_m - r_{f(t+1)}} w_x\right)^2 \hat{\sigma}^2(\bar{r}_m) + w_x^2 \hat{\sigma}^2(\tilde{\epsilon}),$$

where we have used the fact that $\hat{\sigma}^2(\bar{r}_x^p) = (\beta_{xm}^p)^2 \hat{\sigma}^2(\bar{r}_m) + \hat{\sigma}^2(\tilde{\epsilon}_x)$ and $\hat{\text{Cov}}(\bar{r}_x^p, \bar{r}_m) = \hat{\beta}_{xm}^p \hat{\sigma}^2(\bar{r}_m)$. The w_x that minimizes the variance is

$$w_x = \frac{\hat{\alpha}_x}{(\bar{r}_m - r_{f(t+1)}) \frac{\hat{\sigma}^2(\tilde{\epsilon}_x)}{\hat{\sigma}^2(\bar{r}_m)} + \frac{\hat{\alpha}_x^2}{\bar{r}_m - r_{f(t+1)}}}. \quad (39)$$

Note that the sign of w_x is determined by the sign of $\hat{\alpha}_x$. Relations (39) and (38) determine the weights of the portfolio that has the same expected rate of return as the market and has the minimum variance. It is straightforward to verify that this portfolio, say q , has a variance that is strictly less than that of the market portfolio; see Figure 6. Then one can construct portfolios of q and the riskless asset to form a “super efficient portfolio frontier”.

Any passive portfolio formed by a linear combination of the market portfolio and the riskless asset is dominated in the mean-variance sense by a portfolio on the super efficient portfolio frontier.

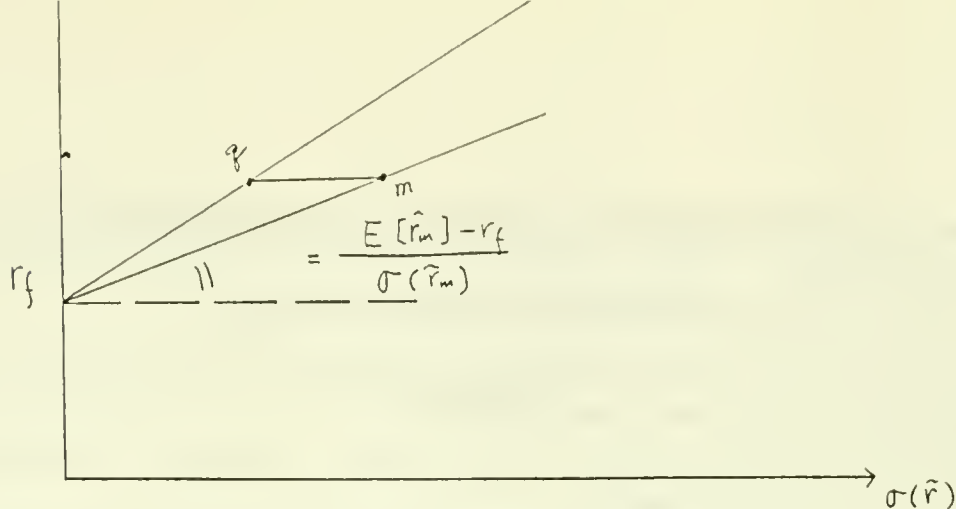


Figure 6: The Super Efficient Portfolio Frontier

An investor can then choose his optimal portfolio by the tangent point between his indifference curves and the super efficient portfolio frontier.

From the earlier analysis, we can also see that a measurement of the the ability of a portfolio manager to select mispriced securities is to see whether his portfolio, say q , has a reward to variability ratio

$$\frac{\bar{r}_q - r_f}{\hat{\sigma}(\bar{r}_q)}$$

that is greater than the slope of the Capital Market Line. More generally, a portfolio is said to outperform another portfolio if the former has a reward to variability ratio greater than that of the latter.

6 Concluding Remarks

In this paper we have briefly reviewed the modern portfolio theory and the Capital Asset Pricing Model. Our review is of course nowhere complete. Readers interested in these subjects are encouraged to consult Brealey and Myers (1988), Huang and Litzenberger (1988), Merton (1982), and Sharpe and Alexander (1990).

7 References

1. Black, F. 1972. Capital market equilibrium with restricted borrowing. *Journal of Business* 45:444-454.
2. Brealey, R. and S. Myers. 1988. *Principles of Corporate Finance*. Third Ed. McGraw Hill. New York.
3. Huang, C. and H. Litzenberger, 1988, *Foundations for Financial Economics*, North-Holland, Amsterdam.
4. Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47:13-37.

5. Lintner, J. 1969. The aggregation of investor's diverse judgements and preferences in purely competitive markets. *Journal of Financial and Quantitative Analysis* 4:346-382.
6. Markowitz, H. 1952. Portfolio Selection. *Journal of Finance* 7:77-91.
7. Merton, R. 1972. An analytical derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis* 7:1851-1872.
8. Merton, R. 1982. *Finance Theory*. Lecture Notes, Sloan School of Management, Massachusetts Institutes of Technology.
9. Mossin, J. 1966. Equilibrium in a capital asset market. *Econometrica* 35:768-783.
10. Myers, S. and S. Turnbull. 1977. Capital Budgeting and the Capital Asset Pricing Model: Good News and Bad News. *Journal of Finance* 32:321-332.
11. Roll, R. 1977. A critique of the asset pricing theory's tests. *Journal of Financial Economics* 4:129-176.
12. Sharpe, W. 1964. Capital asset prices: A theory of capital market equilibrium under conditions of risk. *Journal of Finance* 19:425-442.
13. Sharpe, W. and G. Alexander. 1990. *Investments*. Fourth Ed. Prentice Hall, New Jersey, USA.

Contingent Securities Valuation in Continuous Time

Chi-fu Huang

1 Introduction

The understanding of the pricing of securities whose payoffs are contingent on the prices of other securities is perhaps the major achievement of financial economics in the past two decades. The seminal papers are Black and Scholes (1973) and Merton (1973), which deal mainly with the pricing of options. Subsequent to these two seminal papers, key observations due originally to Cox and Ross (1976) and later formalized by Harrison and Kreps (1979) have made possible a general theory of contingent security valuation.

This general theory of contingent security valuation, although very abstract, turns out to have profound implications in financial engineering. For a practitioner, one does not need to know the intricate workings of the theory. One just has to verify some regularity conditions and follow several mechanical steps to value a contingent security. As a by-product of the valuation procedure, one can also compute a dynamic trading strategy that “replicates” the payoffs of a contingent security.

Rather than developing the general theory here, which is way beyond the scope of the current paper, we will try to introduce the main ideas behind it. The model we use for this purpose is the model originally due to Samuelson (1965) and Black and Scholes (1973), where a geometric Brownian motion is used to represent the price dynamics of a risky security and the riskless interest rate is a constant. We will show that any security whose payoffs depend on the historical prices of the risky security has a well-defined value. This will be done by demonstrating that such payoffs can be “replicated” by a dynamic trading strategy in the risky security and in the riskless security. In particular, we will calculate the value of a European call option written on the risky security to arrive at the Black-Scholes option pricing formula. We will also demonstrate how one finds the replicating strategy for the call option.

Even for this simple model, we cannot hope to make a presentation that is complete with technical details. We will at times simply state the formal results without proofs and refer the reader to the relevant references.

The rest of this paper is organized as follows. Section 2 formulates our model and shows that all contingent securities have well-defined values. In Section 3, the Black-Scholes option pricing formula is discussed and the mechanics of the valuation procedure is recorded. Section 4 provides some generalizations of the results of earlier sections and Section 5 contains some concluding remarks.

2 Formulation

Consider a securities market that has one risky and one riskless asset. We will only be interested in contingent securities having payoffs only at time $T > 0$ that depend upon the historical realizations of the prices of the risky security. Assume that on the time interval $[0, T]$ the risky security does not pay dividends and its price process $S = \{S(t); t \in [0, T]\}$ follows a geometric Brownian motion:

$$S(t) = \exp \left\{ \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma w(t) \right\}, \quad t \in [0, T], \quad (40)$$

where $w = \{w(t); t \in [0, T]\}$ is a standard Brownian motion under a probability P .² Using Itô's lemma, we can express (40) in differential form as

$$dS(t) = \mu S(t) dt + \sigma S(t) dw(t) \quad t \in [0, T], \quad S(0) = 1. \quad (41)$$

Thus μ is the instantaneous expected rate of return on the risky security and σ is the standard deviation on the instantaneous rate of return on the risky security. The price at time t of the riskless security is

$$B(t) = \exp \{rt\} \quad t \in [0, T].$$

Equivalently we can write

$$dB(t) = rB(t)dt \quad t \in [0, T], \quad B(0) = 1.$$

The riskless security can be thought of as a savings account that pays a constant instantaneous interest rate r . Thus if one invests one dollar in the riskless security at time t , then the value of ones investment at time $s \geq t$ is $\exp\{r(s - t)\}$. These two securities will hence be called *long-lived securities* as they are available for trading all the time in $[0, T]$.

If a statement holds with probability one with respect to P we will say that the statement holds P -almost surely, or simply P -a.s.

A contingent security is represented by a random variable x whose value depends upon the historical realizations of S from time 0 to time T . We interpret x to be the payoff of a contingent security, which is delivered at time T . For technical reasons, we will focus our attention to those contingent securities whose payoffs x have a finite second moment; that is, $E[x^2] < \infty$, where $E[\cdot]$ is the expectation under P . Denote this space of contingent security by $L^2(P)$.

Investors in the market can trade in the two long-lived securities any time $t \in [0, T]$. Let $\alpha(t)$ and $\theta(t)$ be the numbers of shares of the riskless security and the risky security held at time t . A natural informational constraint on $\alpha(t)$ and $\theta(t)$ is that their values can only depend upon the realizations of S from time 0 to time t . A trading strategy is said to be *self-financing* if for all $t \in [0, T]$,

$$\alpha(t)B(t) + \theta(t)S(t) = \alpha(0)B(0) + \theta(0)S(0) + \int_0^t \alpha(s)dB(s) + \int_0^t \theta(s)dS(s) \quad P - a.s. \quad (42)$$

That is, the value of the portfolio the trading strategy holds at time t is equal to the initial cost of the strategy plus the accumulated capital gains/losses from trading. After the initial investment,

²A standard Brownian motion w under P is a process having continuous sample paths, stationary independent increments, and $w(t)$ is a Normally distributed random variable with mean zero and variance t under P .

there is neither additional fund invested into nor funds withdrawn out of the strategy and thus the name of the trading strategy.

Note that implicit in the above definition of a self-financing trading strategy is the hypothesis that the integrals of (42) are well-defined. We refer the readers to Liptser and Shiriyayev (1977, chapter 4), for example, for the definition of an integral involving a Brownian motion or an *Itô* integral.

A contingent security $x \in L^2(P)$ is said to be *marketed* if there exists a self-financing trading strategy (α, θ) so that $x = \alpha(T)B(T) + \theta(T)S(T)$ P -a.s. In such event, we will often say x is *financed* by (α, θ) . In words, a contingent security is marketed if one can find a self-financing trading strategy so that the value of the strategy at time T is equal to x with probability one, or x can be “replicated” by dynamic trading in the two long-lived securities. It should be clear that if a contingent security is financed by two different self-financing strategies, the initial investments of these two strategies must be the same if there are no arbitrage opportunities (to be defined formally shortly). Then the value of the contingent security at any time t is well-defined and is equal to the value of its replicating portfolio at that time.

Our task here is to determine which contingent securities are marketed and what are the “fair” prices for them. Before we proceed, however, we have to make sure that the model we have written down is reasonable in that something cannot be created from nothing, or no *arbitrage opportunities* are present, so that marketed contingent securities have well-defined prices. Formally, an arbitrage opportunity is a contingent security $x \in L^2(P)$ that is financed by a self-financing trading strategy (α, θ) so that $x \geq 0$ P -a.s. and $x \neq 0$, and $\alpha(0)B(0) + \theta(0)S(0) \leq 0$. In other words, an arbitrage opportunity is a self-financing trading strategy that requires zero or negative initial investment but can produce a payoff that is positive and nonzero. If such a trading strategy exists, the price processes (B, S) can never occur in any economic equilibrium as an investor will take an infinite position in the strategy and the market will not clear.

It is known that the existence of a probability measure Q that has the same probability zero sets as the probability P does and under which the normalized price process $S^* = \{S^*(t) = S(t)/B(t); t \in [0, T]\}$ is a martingale is a necessary and sufficient condition for there not to be arbitrage opportunities in a discrete time discrete state model; see Huang and Litzenberger (1988, chapter 8) and see the example in Huang (1990a, section 4). Two probability measures having the same probability zero sets are said to be *equivalent* probabilities. Thus Q will be termed an *equivalent martingale measure*. (Note in Huang (1990a) that the word “equivalent” was not mentioned for brevity. The martingale measure constructed there was indeed equivalent, however.)

In a continuous time model, (under some regularity conditions) the existence of an equivalent martingale measure is still a necessary condition for the absence of arbitrage opportunities. In our model, we now demonstrate that this necessary condition is met. We will accomplish this by construction.

Put

$$\eta(t) = \exp \left\{ -\frac{1}{2} \left(\frac{\mu - r}{\sigma} \right)^2 t - \left(\frac{\mu - r}{\sigma} \right) w(t) \right\} \quad t \in [0, T].$$

Using the fact that $w(t)$ is normally distributed and thus $\eta(t)$ is lognormally distributed, one quickly

verifies that $E[\eta(t)] = 1$ for all $t \in [0, T]$. Hence we can define a probability Q by

$$Q(A) = \int_A \eta(\omega, T) P(d\omega)$$

for all events A that are defined by the historical realizations of S from $t = 0$ to $t = T$, where ω denotes one of these possible realizations.

We prove below that Q is an *equivalent martingale measure*. First, as $\eta(T)$ is lognormally distributed, it is strictly positive P -a.s. This immediately implies that Q is equivalent to P . Since P and Q have the same probability zero sets, we will hence simply use a.s. to denote almost surely with respect to either probabilities.

Before proceeding, we record a mathematical result, the celebrated Girsanov theorem:

Lemma 2.1 *Suppose that*

$$\zeta(T) = \exp \left\{ -\frac{1}{2} \int_0^T \beta^2(s) ds + \int_0^T \beta(s) dw(s) \right\}$$

for some process β with $\int_0^T |\beta(t)|^2 dt < \infty$ P -a.s. and that $E[\zeta(T)] = 1$. Then the process

$$\hat{w}(t) = w(t) - \int_0^t \beta(s) dw(s) \quad t \in [0, T] \quad (43)$$

is a standard Brownian motion under the probability

$$\hat{Q}(A) = \int_A \zeta(\omega, T) P(d\omega)$$

for all events A that are defined by the historical realizations of S from $t = 0$ to $t = T$.

PROOF. See, for example, Liptser and Shirayev (1977, 6.3). ■

From Girsanov theorem, we know that under Q

$$w^*(t) \equiv w(t) + \frac{\mu - r}{\sigma} t, \quad t \in [0, T] \quad (44)$$

is a standard Brownian motion. Next note that Itô's lemma allows us to write

$$\begin{aligned} dS^*(t) = d(S(t)/B(t)) &= (\mu - r)S^*(t)dt + \sigma S^*(t)dw(t) \\ &= (\mu - r)S^*(t)dt + \sigma S^*(t) \left(dw^*(t) - \frac{\mu - r}{\sigma} dt \right) \\ &= \sigma S^*(t)dw^*(t). \end{aligned} \quad (45)$$

Or equivalently,

$$S^*(s) = S^*(t) \exp \left\{ -\frac{1}{2} \sigma^2 (s - t) + \sigma (w^*(s) - w^*(t)) \right\}, \quad s \geq t. \quad (46)$$

Taking the conditional expectation of $S^*(s)$ under Q at time t gives

$$E^*[S^*(s)|\mathcal{F}_t] = S^*(t) \quad Q - a.s.,$$

where $\mathcal{F}_t$ denotes the information one has at time t by observing the prices of the risky long-lived security from time 0 to time t and $E^*[\cdot|\mathcal{F}_t]$ is the conditional expectations operator under Q given $\mathcal{F}_t$. This shows that S^* is a martingale under Q and Q is an equivalent martingale measure. Technical arguments can further show that Q is actually the unique equivalent martingale measure since the Brownian motion is one dimensional and there is one risky security.

Using (41) and (44), we can write

$$dS(t) = rS(t)dt + \sigma S(t)dw^*(t), \quad t \in [0, T], \quad S(0) = 1. \quad (47)$$

That is, under the equivalent martingale measure Q , the instantaneous expected rate of return on the long-lived risky security is equal to the riskless interest rate. For this reason, Q is also called the *risk neutral probability* as in a world where all the investors are risk neutral, in equilibrium, it is necessary that all securities, risky as well as riskless, must be making the same rate of return.

Without further restriction on the trading strategies other than the requirement that the Itô integral of (42) is well defined, it is known that arbitrage opportunities exist. A strategy that mimics the doubling of one's bet in a roulette has been shown by Harrison and Kreps (1979) to create an arbitrage opportunity. A discussion of the proper restriction on the trading strategies is outside the scope of this paper, for which we refer the reader to Dybvig and Huang (1989) and Pages (1989). Here we simply say we will restrict our attention to those self-financing trading strategies (α, θ) so that

$$E \left[\int_0^T |\theta(t)S(t)|^2 dt \right] < \infty.$$

Denote this space of trading strategies by $H^2(P)$.

We now show that no arbitrage opportunities can be created using strategies from $H^2(P)$. First note that if $x \in L^2(P)$ is financed by a self-financing trading strategy (α, θ) , then the value of the strategy at time t in units of the riskless security is $\alpha(t) + \theta(t)S^*(t)$, which by Itô's lemma can be expressed as

$$\alpha(t) + \theta(t)S^*(t) = \alpha(0) + \theta(0)S^*(0) + \int_0^t \theta(s)dS^*(s). \quad (48)$$

Pages (1989) shows that if $(\alpha, \theta) \in H^2(P)$, then the Itô integral on the right-hand side of (48) is a martingale under Q . Thus taking expectations under Q on both sides gives

$$E^*[\alpha(t) + \theta(t)S^*(t)] = \alpha(0) + \theta(0)S^*(0). \quad (49)$$

This is the first main result of this paper. The expectation under Q is the "present value operator" that gives the time zero value, in units of the riskless security, of a random payoff delivered at any time after time zero. In particular, since $\alpha(T) + \theta(T)S^*(T) = xe^{-rT}$, (49) implies that

$$E^*[xe^{-rT}] = \alpha(0) + \theta(0)S^*(0).$$

That is, the value at time zero of a marketed contingent security, in units of the riskless security, is equal to the expectation under Q of the contingent security's payoff at T in units of the riskless security.

Similar arguments show that

$$\alpha(t) + \theta(t)S^*(t) = E^* \left[xe^{-rT} | \mathcal{F}_t \right]. \quad (50)$$

That is, the value over time of a marketed contingent security, in units of the riskless security, is a martingale under the equivalent martingale measure.

Now we are ready to show that $H^2(P)$ does not contain any arbitrage opportunities. Suppose on the contrary that there is $(\alpha, \theta) \in H^2(P)$ with $\alpha(0)B(0) + \theta(0)S(0) \leq 0$ and $\alpha(T)B(T) + \theta(T)S(T) \geq 0$ a.s. and $\alpha(T)B(T) + \theta(T)S(T) \neq 0$. By (49) we know that

$$\alpha(0)B(0) + \theta(0)S(0) = B(0)E^*[\alpha(T) + \theta(T)S^*(T)].$$

Since $\alpha(T) + \theta(T)S^*(T) = (\alpha(T)B(T) + \theta(T)S(T))e^{-rT}$ is positive and nonzero respect to both P and Q by hypothesis, the right-hand side of the above relation is strictly positive. This contradicts the hypothesis that the left-hand side is less than or equal to zero. Hence there are no arbitrage opportunities.

So the lesson to remember so far is that our model is a reasonable one in that there are no arbitrage opportunities and thus marketed contingent securities have well-defined prices. More important, in the model, in units of the riskless security, not only the prices over time of the risky long-lived security but also those of the marketed contingent securities, form a martingale under Q .

At first sight, the above martingale connection of the prices of marketed contingent securities does not seem to be very helpful if one can identify a marketed contingent security only by first demonstrating its replicating strategy. Thus the martingale connection does not help us calculate the prices over time of a marketed contingent security as we know those prices already once we know the replicating strategy.

What we will show next is that some abstract arguments can be made to show that *all* the contingent securities in $L^2(P)$ are marketed without explicitly demonstrating each and every replicating strategy. Then the martingale connection becomes extremely useful as we can then calculate the values over time of each and every contingent security simply by evaluating conditional expectations. The reader will later find out that in many situations, the ability to calculate the values over time of a contingent security also provides a way to construct the contingent security's replicating strategy.

First note that if we can show for any $x \in L^2(P)$ there exists a process θ with

$$E \left[\int_0^T |\theta(t)S(t)|^2 dt \right] < \infty$$

so that for all t ,

$$E^*[xe^{-rT} | \mathcal{F}_t] = E^*[xe^{-rT}] + \int_0^t \theta(s)dS^*(s), \quad (51)$$

then by defining

$$\alpha(t) = E^*[xe^{-rT} | \mathcal{F}_t] - \theta(t)S^*(t), \quad (52)$$

one can show using Itô's lemma that x is financed by $(\alpha, \theta) \in H^2(P)$. Using a martingale representation theorem due originally to Kunita and Watanabe (1965), Pages (1989) did just that. Thus all contingent securities are marketed and their prices over time are determined according to (50).

3 Black-Scholes Option Pricing Formula

We turn in this section our attention to some applications of the model developed in the previous section. In particular, we will calculate the prices over time of a European call option written on the risky long-lived security with a maturity date T and an exercise price K . This is a financial contract that gives its owner the right to buy one share of the risky long-lived security for a price K at time T . At the maturity date of this option, one will exercise the right to purchase if and only if $S(T) \geq K$. Otherwise, the option will expire worthless. Thus the payoff at time T of this contract is

$$\max[S(T) - K, 0].$$

From the discussion in Section 2, we know that this call option is marketed and whose price at time t , $C(t)$, by (50) is

$$C(t) \equiv e^{rt} E^*[\max[S(T) - K, 0]e^{-rT} | \mathcal{F}_t] = e^{rt} E^*[\max[S^*(T) - Ke^{-rT}, 0] | \mathcal{F}_t].$$

By the fact that a Brownian motion has independent increments, the distribution of $S^*(T)$ conditional $\mathcal{F}_t$ only depends on $S^*(t)$. Recall then from (46) that under Q , $S^*(T)$ is lognormally distributed conditional on $S^*(t)$. We can thus evaluate the conditional expectation explicitly to get

$$C(t) = S(t)N(d_1) - Ke^{-r(T-t)}N(d_1 - \sigma\sqrt{T-t}), \quad (53)$$

where

$$d_1 = \frac{\ln(S(t)/K) + r(T-t) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}},$$

and $N(\cdot)$ denotes the cumulative distribution function of a standard Normal random variable. This is the option pricing formula due originally to Black and Scholes (1973) and developed further by Merton (1973).

Besides the parameters of the model such as σ , r , and T , the option price at time t depends on the price of its underlying asset $S(t)$, the exercise price K , and the time t . It does not depend on μ .³ For brevity of notation, we will suppress the dependence of $C(t)$ on σ , r , T and K , and write

$$C(t) = C(S(t), t).$$

Next we will show that the trading strategy that finances the call option can actually be constructed from $C(S(t), t)$ by differentiation. For this, we will use the martingale property of $C(t)e^{-rt}$. Since $C(S(t), t)$ is twice continuously differentiable in $S(t)$ and continuously differentiable in t when $t < T$, Itô's lemma allows us to write

$$\begin{aligned} & d(C(S(t), t)e^{-rt}) \\ &= \overbrace{\left[\frac{1}{2}e^{-rt}C_{SS}(S(t), t)S^2(t)\sigma^2 + re^{-rt}C_S(S(t), t)S(t) - re^{-rt}C(S(t), t) + e^{-rt}C_t(S(t), t) \right]}^{\text{time trend}} dt \\ & \quad + e^{-rt}C_S(S(t), t)S(t)\sigma dw^*(t), \end{aligned}$$

³The reason that the option price is independent of μ is outside the scope of this paper. The interested reader is referred to Huang (1989) for a discussion.

where we have used C_S and C_{SS} to denote the partial and the second partial derivative, respectively, of C with respect to S , and C_t to denote the partial derivative of C with respect to t . Since $C(S(t), t)e^{-rt}$ is a martingale under Q , it must not have any time trend. Thus it must be that

$$0 = e^{-rt} \left(\frac{1}{2} C_{SS}(S(t), t) S^2(t) \sigma^2 + r C_S(S(t), t) S(t) - r C(S(t), t) + C_t(S(t), t) \right), \quad (54)$$

for all possible $S(t)$ and $t < T$. We will come back to this equation a little later. For now, (54) implies that

$$\begin{aligned} C(S(t), t) e^{-rt} &= C(S(0), 0) + \int_0^t e^{-rs} C_S(S(s), s) S(s) \sigma dw^*(s) \\ &= C(S(0), 0) + \int_0^t C_S(S(s), s) dS^*(s) \quad \text{a.s. } t \in [0, T]. \end{aligned}$$

Comparing this relation with (51), we let, for $t < T$, $\theta(t) \equiv C_S(S(t), t)$ and, for $t = T$,

$$\theta(T) = \begin{cases} 1 & \text{if } S(T) \geq K; \\ 0 & \text{if } S(T) < K. \end{cases}$$

Direct computation yields $C_S(S(t), t) = N(d_1)$. Then define $\alpha(t)$ according to (52) to have

$$\alpha(t) = (C(S(t), t) - \theta(t)S(t))e^{-rt} = -Ke^{-rT}N(d_1 - \sigma\sqrt{T-t}) \quad \forall t \in [0, T].$$

It is obvious that $(\alpha, \theta) \in H^2(P)$ as $\theta(t) \leq 1$ and finances the European call option. We can reconfirm this calculation by observing the Black-Scholes formula of (53) that the value of the call is exactly equal to the value of its replicating portfolio, which is just

$$\alpha(t)B(t) + \theta(t)S(t) = -Ke^{-r(T-t)}N(d_1 - \sigma\sqrt{T-t}) + S(t)N(d_1).$$

This relation also tells us that a call option is equivalent to a levered position of its underlying security and the leverage changes over time.

Finally, observe that since (54) holds for all $S(t)$ and for all $t \in [0, T]$, $C(\cdot, \cdot)$ satisfies the partial differential equation

$$\frac{1}{2} C_{xx}(x, t) x^2 \sigma^2 + r x C_x(x, t) - r C(x, t) + C_t(x, t) = 0 \quad (55)$$

for all $x \in (0, \infty)$ and $t \in [0, T]$, with the boundary condition

$$\begin{aligned} \lim_{t \uparrow T} C(x, t) &= \max[x - K, 0], \\ \lim_{x \downarrow 0} C(x, t) &= 0 \quad \forall t \in [0, T]. \end{aligned}$$

Note that in deriving the differential equation of (55), the only thing we made use of is the fact that $C(t)$ is a smooth function of $S(t)$ and t and that $C(t)e^{-rt}$ is a martingale under Q . Therefore, this equation should hold more generally. Once we show that the price at time t of a contingent

security is a smooth function of $S(t)$ and t , then it follows that the price functional of this contingent security must satisfy (55) by Itô's lemma and the martingale connection. The only thing changed is the boundary conditions, which reflect the contractual agreement of a contingent security.

For example, consider a European put option that gives its owner the right to sell a share of the risky long-lived security for a price K at time T . So the payoffs at time T is $\max[K - S(T), 0]$. The price of this put option at time t is

$$\begin{aligned} p(t) &= E^*[\max[K - S(T), 0]e^{-r(T-t)}|\mathcal{F}_t] \\ &= Ke^{-r(T-t)}N(\hat{d}_1) - SN(\hat{d}_1 - \sigma\sqrt{T-t}), \end{aligned}$$

where

$$\hat{d}_1 = \frac{\ln(K/S(t)) - r(T-t) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}.$$

Since $p(t)$ is a smooth function of $S(t)$ and t , $p(S(t), t)$ must satisfy (55) with the following boundary conditions

$$\begin{aligned} \lim_{t \uparrow T} p(x, t) &= \max[K - x, 0], \\ \lim_{x \downarrow \infty} p(x, t) &= 0 \quad \forall t \in [0, T). \end{aligned}$$

To determine the replicating strategy for the put, we simply set $\theta(t) = p_S(S(t), t)$ and define $\alpha(t)$ according to (52).

In this case, again, we do not make use of the the differential equation. In the following example, however, one could not evaluate the conditional expectation of (50) analytically and thus the differential equation becomes useful. One can numerically solve the differential equation to get the price of the contingent security.

Consider now a contingent security that pays $[\sin(S(T))]^3$ at time T . It is easy to see that the conditional expectation

$$E^*[e^{-r(T-t)}[\sin S(T)]^3|\mathcal{F}_t]$$

depends only on $S(t)$ and t even though we cannot evaluate it analytically. Given that the payoff is a smooth function of $S(T)$, one can also show that the conditional expectation must also be a smooth function of $S(t)$ and t . Let this function be $g(S(t), t)$. Then g must satisfy the partial differential equation of (55) with the boundary conditions that $\lim_{t \uparrow T} g(x, t) = (\sin x)^3$ and $\lim_{x \downarrow 0} g(x, t) = 0$. One can compute g by solving the differential equation. Of course, one can also numerically compute the conditional expectation by using numerical integration or simulation. The choice of the techniques will depend upon their relative efficiencies and is outside the scope of our discussion.

In any case, once one gets the values of the function g , the replicating strategy can easily be constructed by taking the derivative of g and setting $\theta(t) = g_S(S(t), t)$. Then use (52) to calculate α .

4 Generalizations

In our discussion in Section 3, we presented how one could construct the replicating strategies for contingent securities whose payoffs at time T are a function of $S(T)$. We also discussed the

usefulness of the differential equation of (55). Here we will first show how one can generalize all this to contingent securities whose payoffs at time T depend on the historical prices of the long-lived risky security. Take for example a contingent security that gives its owner the right to receive an exponentially weighted average of the past prices of the long-lived risky security for a fixed fee of K . Formally, let this weighted average be

$$z(T) = \beta \int_0^T e^{-\beta(T-t)} S(t) dt$$

and the payoffs of this contingent security is then

$$\max[z(T) - K, 0].$$

The value at time t of this contingent security is

$$E^*[\max[z(T) - K, 0] e^{-r(T-t)} | \mathcal{F}_t], \quad (56)$$

which is by now mechanical. The distribution under which we evaluate this conditional expectation is the probability under which the risky long-lived security earns an expected rate of return equal to the riskless rate. We show next that this conditional expectation can be calculated by solving a differential equation which is slightly different from (55).

Put

$$z(t) = \beta \int_0^t e^{-\beta(t-s)} S(s) ds;$$

or in differential form:

$$dz(t) = \beta(S(t) - z(t))dt, \quad z(0) = 0.$$

Let $s > t$. We can write

$$z(s) = z(t)e^{-\beta(s-t)} + \beta \int_t^s e^{-\beta(s-\tau)} S(\tau) d\tau.$$

That is, knowing $z(t)$, the distribution of $z(s)$ only depends on S from t to s , whose distribution in turn only depends upon $S(t)$. Hence (56) can be written as a function of $S(t)$, $z(t)$, and t . Let this function be $h(S(t), z(t), t)$. Argument can be made to show that h is smooth in its arguments.

Next use Itô's lemma to express $h(S(t), z(t), t)e^{-rt}$ to get

$$\begin{aligned} & d(h(S(t), z(t), t)e^{-rt}) \\ = & \left[\frac{1}{2} e^{-rt} h_{SS}(S(t), z(t), t) S^2(t) \sigma^2 + r e^{-rt} h_S(S(t), z(t), t) S(t) \right. \\ & \quad \left. + e^{-rt} h_z(S(t), z(t), t) \beta(S(t) - z(t)) - r e^{-rt} h(S(t), z(t), t) + e^{-rt} h_t(S(t), z(t), t) \right] dt \\ & \quad + e^{-rt} h_S(S(t), z(t), t) S(t) \sigma dw^*(t). \end{aligned}$$

Since $h(S(t), z(t), t)e^{-rt}$ is a martingale under Q and cannot have any time trend, we must have

$$\frac{1}{2} h_{SS}(S, z, t) S^2 \sigma^2 + r h_S(S, z, t) S + h_z(S, z, t) \beta(S - z) - r h(S, z, t) + h_t(S, z, t) = 0 \quad (57)$$

for all S and z . This is a differential equation that h must satisfy. The boundary conditions are

$$\begin{aligned}\lim_{t \uparrow T} h(S, z, t) &= \max[z - K, 0], \\ \lim_{S \downarrow 0} h(S, z, t) &= \max[ze^{-\beta(T-t)} - K, 0]e^{-r(T-t)} \quad \forall t \in [0, T].\end{aligned}$$

The replicating strategy can then be mechanically constructed. Define $\theta(t) = h_S(S(t), z(t), t)$ and $\alpha(t)$ according to (52). The replicating strategy at time t now depends not only on $S(t)$ and t but also on $z(t)$.

The methodology presented in previous sections also generalizes to economies with more general market structures. For example, there can be an arbitrary number of risky long-lived securities whose price processes are general Itô processes and the riskless interest rate can be a random process too. As long as the number of linearly independent risky long-lived securities is equal to the number of the independent Brownian motions that drive the price processes, under some regularity conditions, there exists a unique equivalent martingale measure and all the contingent securities in $L^2(P)$ are marketed. This unique martingale measure is also a *risk-neutral probability* that makes all the securities earn an expected rate of return equal to the riskless interest rate. (Compare with (47).)

Letting $E^*[\cdot | \mathcal{F}_t]$ denote the conditional expectation operator at time t , the price at time t of a contingent security with a payoff x is

$$E^* \left[x e^{-\int_t^T r(s) ds} | \mathcal{F}_t \right].$$

This conditional expectation can be calculated in a variety of ways when its analytic expression does not exist. When it can be identified to be a smooth function of several variables, it satisfies partial differential equation through its connection to martingales. The boundary conditions will be determined by the contractual agreements. In this case, the replicating strategy can be computed by taking derivatives of this function. When the conditional expectation cannot be expressed as a function of a finite number of variables, it can always be approximated by using simulation.

5 Concluding Remarks

We have tried in this paper to introduce the main ideas behind perhaps the most important advancement in financial economics in the past two decades. Our discussion in Section 2 is essentially abstract in nature. We do hope, however, that the applications presented in Sections 3 and 4 convey to the reader the engineering aspect of the abstract theory: Once the unique equivalent martingale measure is identified, the values over time of a contingent security can be computed by evaluating a conditional expectation. When this conditional expectation is a function of several variables, this function satisfies a partial differential equation. In such case, the replicating strategy for the contingent security can be calculated by taking derivatives of this function. So, the procedure in application becomes mechanical.

The reader interested in detailed discussions of the general theory of contingent securities valuation is referred to Duffie (1988) and Huang (1989).

6 References

1. Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81:637–654.
2. Cox, J., S. Ross. 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3:145–166.
3. Duffie, D., 1988, *Security Markets: Stochastic Models*, Academic Press,
4. Dybvig, D., and C. Huang, 1989, Nonnegative Wealth, Absence of Arbitrage and Feasible Consumption Plans, *Review of Financial Studies* 1, 377–401.
5. Harrison, M., and D. Kreps. 1979. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* 20:381–408.
6. Harrison, M., and S. Pliska. 1981. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications* 11:215–260.
7. Huang, C., and R. Litzenberger, 1988, *Foundations for Financial Economics*, North-Holland, New York.
8. Huang, C., 1990a, An overview of modern financial economics, the first article in this collection.
9. Huang, C., 1989, *Lecture Notes on Advanced Financial Economics*, Sloan School of Management, Massachusetts Institute of Technology.
10. Kunita, H., and S. Watanabe, 1967, On square-integrable martingales, *Nagoya Mathematics Journal* 30:209–245.
11. Liptser, R., and A. Shiriyayev, 1977, *Statistics of Random Processes I: General Theory*, Springer-Verlag, New York.
12. Merton, R. 1973. Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4:141–183.
13. Pagès, H. 1989. Three Essays In Optimal Consumption. Unpublished Ph.D. thesis. Massachusetts Institutes of Technology.
14. Samuelson, P., 1965, Rational theory of warrant pricing, *Industrial Management Review* 6:13–32.

Dynamic Portfolio Theory and the Intertemporal Capital Asset Pricing Model

Chi-fu Huang

1 Introduction

Traditionally, the optimal consumption and portfolio decisions of an investor in multiperiod economies under uncertainty, or the *dynamic portfolio theory*, is developed using stochastic dynamic programming. Merton (1971) is the seminal paper in this area. In this paper, however, instead of using the traditional approach, we will outline how the theory of the contingent securities valuation summarized in Huang (1990c) can be useful in analyzing an investor's optimal consumption and portfolio policies. We will show in particular that the optimal policy of an investor can be computed by solving a linear partial differential equation similar to the one appears in the valuation of contingent securities.

Once we have a characterization of an investor's optimal consumption and portfolio policies, it is then an easy matter to draw conclusions about the equilibrium relation among the expected rates of return on securities in an economy with a single representative investor. This is the *Intertemporal Capital security Pricing Model* of Merton (1973).

2 Optimal Consumption and Portfolio Policies

We begin our discussion of the optimal consumption and portfolio policy for an investor in the Black-Scholes model. Then we will discuss how the results in the Black-Scholes model can be generalized.

2.1 The Black-Scholes Model

The setup here is identical to that in Section 2 of Huang (1990c), to which we will refer the reader for details.

There are two long-lived securities, one risky and one riskless. The risky security price follows a geometric Brownian motion

$$S(t) = \exp \left\{ \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma w(t) \right\},$$

or in differential form

$$dS(t) = \mu S(t)dt + \sigma S(t)dw(t),$$

where w is a standard Brownian motion under a probability P . The price of the riskless security at time t is $B(t) = \exp\{rt\}$, where $r > 0$ is a constant. One can think of the riskless security as a savings account that pays a constant instantaneous interest rate r . The information an investor has at time t is the historical realizations of the risky security prices from time 0 to time t . Denote this information by $\mathcal{F}_t$.

An investor in this economy uses a *self-financing* strategy (α, θ) to manage his portfolio of the risky and the riskless securities so as to maximize the expected utility of his wealth at time T ; see Huang (1990c) for the definition of a self-financing strategy. Recall that $\alpha(t)$ and $\theta(t)$ are the numbers of shares of the riskless and the risky securities, respectively, held at time t . As in Huang (1990c), we will restrict our attention to those self-financing strategies (α, θ) with

$$E \left[\int_0^T |\theta(t)S(t)|^2 dt \right] < \infty.$$

Call these self-financing trading strategies admissible trading strategies and denote them by $H^2(P)$. It is known that there are no arbitrage opportunities in $H^2(P)$; see Huang (1990c, section 3.2).

Letting $L^2(P)$ be the space of random variables whose values depend on $\mathcal{F}_T$, an investor's task is to solve the following program:

$$\begin{aligned} \sup_{(\alpha, \theta) \in H^2(P)} & E[V(W_T)] \\ \text{s.t.} & \alpha(0)B(0) + \theta(0)S(0) = W_0, \\ & W_T \equiv \alpha(T)B(T) + \theta(T)S(T) \in L^2(P), \\ & W_T \geq 0. \quad P - a.s., \end{aligned} \tag{58}$$

where $V(\cdot)$ is the investor's utility function for time T wealth, $E[\cdot]$ is the expectation under P and $W_0 > 0$ is the investor's initial wealth. We assume that V is continuous, increasing, strictly concave, and differentiable.

Recall from Huang (1990c, sections 3.2 and 3.3) that the markets are dynamically complete in the Black-Scholes model in that every $x \in L^2(P)$ is financed by some $(\alpha, \theta) \in H^2(P)$, and the value at time t a contingent security x financed by $(\alpha, \theta) \in H^2(P)$ is

$$E^*[e^{-r(T-t)}x|\mathcal{F}_t], \quad Q - a.s., \tag{59}$$

where $E^*[\cdot|\mathcal{F}_t]$ is the conditional expectation at time t under the unique equivalent martingale measure Q . The unique equivalent martingale measure Q is defined by

$$\begin{aligned} Q(A) &= \int_A \eta(T)P(d\omega), \\ \eta(t) &= \exp \left\{ -\frac{1}{2}\kappa^2 t + \kappa w(t) \right\}, \\ \kappa &\equiv -\frac{\mu - r}{\sigma}, \end{aligned}$$

for all events A in $\mathcal{F}_T$. Since P and Q are two equivalent probabilities, we will simply use *a.s.* to denote "almost surely" with respect to both.

The dynamic completeness of markets and (59) allow us to transform the dynamic program of (58) into an equivalent static program.

Theorem 2.1 *If (α, θ) is a solution to (58), then $W_T^* = \alpha(T)B(T) + \theta(T)S(T)$ a.s. is a solution to*

$$\begin{aligned} \sup_{W_T \in L_+^2(P)} & E[V(W_T)] \\ \text{s.t.} & E^*[e^{-rT}W_T] = W_0. \end{aligned} \quad (60)$$

Conversely, let W_T^* be a solution to (60). Then W_T^* is financed by some $(\alpha, \theta) \in H^2(P)$ and (α, θ) is a solution to (58).

PROOF. Suppose that (α, θ) is a solution to (58). We claim that $W_T^* = \alpha(T)B(T) + \theta(T)S(T) \in L_+^2(P)$ a.s. is a solution to (60). First note that $E^*[e^{-rT}W_T^*] = \alpha(0)B(0) + \theta(0)S(0) = W_0$. So W_T^* is feasible in (60). If W_T^* is not a solution to (60), there must exist $\hat{W}_T$ feasible in (60) so that $E[V(\hat{W}_T)] > E[V(W_T^*)]$. As the markets are dynamically complete, $\hat{W}_T$ is financed by some $(\hat{\alpha}, \hat{\theta}) \in H^2(P)$ with $\hat{\alpha}(0)B(0) + \hat{\theta}(0)S(0) = E^*[e^{-rT}\hat{W}_T]$. So $\hat{W}_T$ is a feasible time T wealth in (58). This contradicts the hypothesis that (α, θ) is a solution to (58).

Conversely, let W_T^* be the solution to (60). Since $W_T^* \in L^2(P)$, Huang (1990c, section 3.2 and theorem 3) implies that there exists $(\alpha, \theta) \in H^2(P)$ that finances it with $\alpha(0)B(0) + \theta(0)S(0) = E^*[e^{-rT}W_T^*] = W_0$. Thus W_T^* is a feasible time T wealth in (58). Suppose that (α, θ) is not a solution to (58). Then there exists $(\hat{\alpha}, \hat{\theta}) \in H^2(P)$ that is feasible in (58) and

$$E[V(\hat{W}_T)] > E[V(W_T^*)],$$

where $\hat{W}_T = \hat{\alpha}(T) + \hat{\theta}(T)S(T)$ P -a.s. Since $\hat{W}_T$ is financed by a trading strategy $(\hat{\alpha}, \hat{\theta}) \in H^2(P)$, we know from Huang (1990c, section 3.2) that

$$E^*[e^{-rT}\hat{W}_T] = \hat{\alpha}(0)B(0) + \hat{\theta}(0)S(0) = W_0.$$

Thus $\hat{W}_T$ is feasible in (60). This contradicts the hypothesis that W_T^* is a solution to (60) and thus (α, θ) is a solution to (58). ■

Theorem 2.1 demonstrates the equivalence between a complicated dynamic optimization problem of (58) and a standard static optimization problem of (60). We will assume henceforth that there exists a solution to (60), denoted by W_T^* , and refer readers to Cox and Huang (1986) for details.

We now proceed to construct an optimal portfolio policy. The reader will find out that this is nothing more than finding the replicating strategy for a contingent security as discussed in Huang (1990c).

First, from Lagrangian theory, if there exists a solution to (60), there must exist a strictly positive constant $\lambda > 0$ so that

$$V'(W_T^*) \leq \lambda e^{-rT} \eta(T) \quad \text{a.s.},$$

with the equality holding when $W_T^* > 0$, where V' denotes the derivative of V and where we have used the fact that

$$E^*[e^{-rT}W_T^*] = E[e^{-rT}W_T^*\eta(T)].$$

Define a process

$$\xi(t) = \lambda e^{-rt} \eta(t), \quad t \in [0, T],$$

and the “inverse” of V' by $f(y) = \inf\{x \geq 0 : V'(x) \leq y\}$. Then we have $W_T^* = f(\xi(T))$. Note that $\xi(0) = \lambda$ as $\eta(0) = 1$ and

$$\xi(t) = \xi(s) \exp \left\{ - \left(r + \frac{1}{2} \kappa^2 \right) (t - s) + \kappa (w(t) - w(s)) \right\} \quad \forall t \geq s.$$

The latter relation implies that the conditional probability at time s of $\xi(t)$ with $t \geq s$ is completely determined by $\xi(s)$ as a standard Brownian motion has independent and stationary increments. The following differential form of the dynamics of ξ under Q will be useful

$$d\xi(t) = -\xi(t)(r - \kappa^2)dt + \xi(t)\kappa dw^*(t) \quad t \in [0, T],$$

where

$$w^*(t) \equiv w(t) - \kappa t \quad t \in [0, T]$$

is a standard Brownian motion under Q ; see Huang (1990c, theorem 1).

The value at time t of the optimal time T wealth is

$$\begin{aligned} E^*[e^{-r(T-t)} W_T^* | \mathcal{F}_t] &= E^*[e^{-r(T-t)} f(\xi(T)) | \mathcal{F}_t] \\ &= \frac{E[e^{-r(T-t)} f(\xi(T)) \eta(T) | \mathcal{F}_t]}{\eta(t)} \\ &= \frac{E[f(\xi(T)) \xi(T) | \xi(t)]}{\xi(t)} \equiv F(\xi(t), t), \end{aligned} \quad (61)$$

where the second equality follows from the definition of Q and the conditional Bayes rule and the third equality follows from the fact that given $\xi(t)$, the conditional probability of $\xi(T)$ is completely determined.

So $F(\xi(t), t)$ is the value at time t of the “contingent security” that pays $f(\xi(T))$ at time T . We show below that F satisfies a partial differential equation by using its connection to martingales. This martingale connection also allows us to construct the replicating strategy for $f(\xi(T))$. Since $F(\xi(t), t)$ is the value at t of a contingent security, $e^{-rt} F(\xi(t), t)$ must be a martingale under the equivalent martingale measure Q . Itô's lemma implies that, under Q ,

$$\begin{aligned} d(e^{-rt} F(\xi(t), t)) &= \overbrace{e^{-rt} \left[\frac{1}{2} F_{\xi\xi} \xi^2 \kappa^2 - F_{\xi} (r - \kappa^2) \xi - rF + F_t \right]}^{\text{time trend}} \\ &\quad + F_{\xi} \xi \kappa \frac{1}{\sigma S(t)} dS^*(t), \end{aligned} \quad (62)$$

where $S^*(t) = S(t)/B(t)$ and we have used Huang (1990c, relation (6)). Since the left-hand-side of (62) is a martingale under Q and S^* is also a martingale under Q , the time trend on the right-hand-side must be zero. Thus we have a partial differential equation:

$$\frac{1}{2} F_{\xi\xi} \xi^2 \kappa^2 - F_{\xi} (r - \kappa^2) \xi - rF + F_t = 0, \quad (63)$$

to which we will return later. Then the integral form of (62) becomes

$$e^{-rt} F(\xi(t), t) = F(\xi(0), 0) + \int_0^t \theta(t) dS^*(t),$$

where

$$\theta(t) = F_\xi \xi \kappa \frac{1}{\sigma S(t)}.$$

Putting $\alpha(t) = (F(\xi(t), t) - \theta(t)S(t))/B(t)$, the same arguments proving Huang (1990c, theorem 3) show that $(\alpha, \theta) \in H^2(P)$ and it finances W_T^* with an initial wealth W_0 .

Note that the replicating strategy for W_T^* at any time t only depends on the value of $\xi(t)$ and t , in addition to the parameters μ , σ , and r , since there exists a one-to-one mapping between $S(t)$ and $\xi(t)$. In the theory of optimal controls, one usually looks for controls that are functions of the controlled process, or *feedback controls*. Here, the process being controlled is the value over time of the portfolio or the investor's wealth over time. We now show that the *optimal controls*, (α, θ) , can be written as functions of the investor's wealth and time.

First, put $W(t) = \alpha(t)B(t) + \theta(t)S(t)$, the value of the optimal portfolio at time t . By the definition of F , we have $W(t) = F(\xi(t), t)$. Given that f is a decreasing function and is not a constant, F is a strictly decreasing function of ξ . Hence $\xi(t) = F^{-1}(W(t), t)$, where $F^{-1}(\cdot, t)$ denotes the inverse of $F(\cdot, t)$. Second, recall that the value of the optimal strategy at any time t is only a function of $\xi(t)$ and t . Now since $\xi(t) = F^{-1}(W(t), t)$, the optimal strategy can be written as a function of the wealth and time. We have thus expressed the optimal trading strategies as *feedback controls*.

In the above construction of the optimal strategy, we implicitly assumed that we knew the Lagrangian multiplier λ , which in application needs to be computed. If we have an analytic expression of the conditional expectation defining F for an arbitrary $\xi(0)$, the matter is easy. This is so because we know that λ must be such that $F(\lambda, 0) = W_0$. Hence define $\lambda = F^{-1}(W_0, 0)$ and set $\xi(0) = \lambda$. Then the process ξ is well-defined and so is the optimal trading policy.

When there is no analytic expression for F , the matter is somewhat more complicated. We will resort to solving numerically the partial differential equation of (63) with the boundary conditions of $F(\xi, T) = f(\xi)$ and $\lim_{\xi \rightarrow \infty} F(\xi, t) = 0$ for all t . Once we have computed numerically the values of $F(\cdot, 0)$ for various possible $\xi(0)$. The Lagrangian multiplier can be found by setting $\lambda \equiv F^{-1}(W_0, 0)$. Then the optimal trading strategy is well-specified.

The analogy between how we construct the optimal trading strategy for an investor and the replicating strategy for a contingent security should be apparent. The optimal time T wealth for an investor is of course a contingent security, whose payoffs are determined by a static maximization program of (60). Then the problem reduces to finding the replicating strategy for this contingent security. To do so, we first calculate the value over time of this contingent security F . The optimal trading strategy is then determined by taking derivatives of F . When F does not have an analytic expression, we solve numerically a linear second order partial differential equation.

There exists one distinct difference in the partial differential equations in this case and the case for the valuation of contingent securities discussed in Huang (1990c), however. The differential equations of (16) and (18) of Huang (1990c) do not depend on the instantaneous expected rate of return on the risky security. On the other hand, (63) depends on μ through κ . This is so because F is a function of ξ and whose dynamics under Q depends on κ . Recall now that there exists, for every t , a one-to-one mapping between $\xi(t)$ and $S(t)$. Let this mapping be denoted $h(S(t), t) = \xi(t)$. The reader can verify that this mapping depends on μ , among other things. Putting $G(S(t), t) = F(h(S(t), t), t)$, we get $G(S(t), t) = F(\xi(t), t)$. Then one easily shows that G satisfies the partial differential equation of (16) of Huang (1990c), which is independent of μ . But

the boundary condition for G at $t = T$ becomes $G(S, T) = f(h(S, T))$, which depends on μ . In this light, one sees that in the pricing of options, the knowledge of μ is not necessary as the payoffs of options do not depend on μ . But we do need to know μ to find the optimal time T wealth for the investor as the optimal time T wealth depends on $\xi(T)$, which in turn depends on μ . If the investor does not know the expected rate of return on the risky security, he won't be able to find the optimal portfolio to maximize his expected utility!

2.2 Generalizations

The way we constructed the optimal trading strategy for an investor in the previous subsection can be extended to economies with more general price processes. Let there be N risky securities and one riskless security. These securities do not pay dividends in $[0, T]$. The n -th risky security sells for $S_n(t)$ at time t . Let $S(t) = (S_1(t), S_2(t), \dots, S_N(t))^T$, where ^{T} denotes "transpose". Assume that

$$S(t) = S(0) + \int_0^t I_S(s) \mu(Y(s), s) ds + \int_0^t I_S(s) \sigma(Y(s), s) dw(s) \quad t \in [0, T],$$

where w is an N -dimensional standard Brownian motion, $I_S(t)$ is an $N \times N$ diagonal matrix with the n -th diagonal element being $S_n(t)$, $\mu(Y, t)$ is an $N \times 1$ vector containing the N instantaneous expected rates of return on the risky securities, $\sigma(Y, t)$ is an $N \times N$ matrix with $\sigma\sigma^T$ being the covariance matrix of the instantaneous returns on the risky securities at time t , Y is an M -dimensional diffusion process with $M \leq N$:

$$Y(t) = Y(0) + \int_0^t \beta(Y(s), s) ds + \int_0^t g(Y(s), s) dw(s) \quad t \in [0, T],$$

and β and g are $M \times 1$ vector and $M \times N$ matrix, respectively. We assume that $\sigma(Y(t), t)$ and $g(Y(t), t)$ are of full rank for all $Y(t)$ and t .

For convenience, we will use $\mu(t)$ and $\sigma(t)$ to denote $\mu(Y(t), t)$ and $\sigma(Y(t), t)$, respectively, and similarly use $\beta(t)$ and $g(t)$ for $\beta(Y(t), t)$ and $g(Y(t), t)$, respectively. The riskless security pays an instantaneous riskless rate of $r(Y(t), t)$, which will sometimes be denoted simply by $r(t)$. The price at time t of the riskless security is $B(t) = \exp\{\int_0^t r(s) ds\}$.

Putting

$$\kappa(Y(t), t) = -\sigma(Y(t), t)^{-1}(\mu(Y(t), t) - r(Y(t), t)),$$

we assume that $|\kappa(t)|^2 \equiv \text{tr}(\sigma(t)\sigma(t)^T)$ is bounded uniformly across time and states of nature. This implies that

$$Q(A) = \int_A \eta(\omega, T) P(d\omega) \quad A \in \mathcal{F}_T,$$

is the unique equivalent martingale measure for the price system, where

$$\eta(t) = \exp\left\{-\frac{1}{2} \int_0^t |\kappa(s)|^2 ds + \int_0^t \kappa(s)^T dw(s)\right\} \quad t \in [0, T];$$

see Cox and Huang (1986), for example.

Let $\alpha(t)$ and $\theta_n(t)$ be the number of shares of the riskless and the n -th risky security held at time t . We will use $\theta(t)$ to denote $(\theta_1(t), \dots, \theta_n(t))^T$. A trading strategy (α, θ) is admissible if it is self-financing and

$$E \left[\int_0^T |\theta(s)^T I_S(s) \sigma(s)|^2 ds \right] < \infty.$$

Let $H^2(P)$ denote the space of admissible trading strategies. A contingent security is an element of $L^2(P)$.

Pagès (1989) has shown that markets are dynamically complete in that any element of $L^2(P)$ is financed by some $(\alpha, \theta) \in H^2(P)$. Given this, Theorem 2.1 is valid while replacing the constraint of (60) by

$$E^* \left[e^{-\int_0^T r(t) dt} W_T^* \right] = W_0,$$

and we assume henceforth that there exists a solution to (60). As in the previous section, let W_T^* be the optimal time T wealth in (60) and define $\xi(t) = \lambda e^{-\int_0^t r(s) ds} \eta(t)$. We have

$$\xi(t) = \xi(s) \exp \left\{ - \int_s^t \left(r(\tau) + \frac{1}{2} |\kappa(Y(\tau), \tau)|^2 \right) d\tau + \int_s^t \kappa(Y(\tau), \tau)^T dw(\tau) \right\} \quad \forall t \geq s. \quad (64)$$

Note that, unlike in the Black-Scholes model, there is no one-to-one relation between $\xi(t)$ and $S(t)$ at every t in our current setup. Indeed, the value of $\xi(t)$ depends on the realizations of Y from time 0 to time t . However, given that Y is a diffusion process and (64), the probabilistic characteristics of $\xi(t)$ at time s with $s \leq t$ is completely determined by $\xi(s)$ and $Y(s)$. In mathematical terms, (ξ, Y) jointly form a diffusion process.

To find the replicating trading strategy for W_T^* , we will calculate the value over time for W_T^* . Since W_T^* is a contingent security, its value at time t is

$$\begin{aligned} E^*[e^{-\int_0^T r(t) dt} W_T^* | \mathcal{F}_t] &= E^*[e^{-\int_t^T r(s) ds} f(\xi(T)) | \mathcal{F}_t] \\ &= \frac{E[e^{-\int_t^T r(s) ds} f(\xi(T)) \eta(T) | \mathcal{F}_t]}{\eta(t)} \\ &= \frac{E[f(\xi(T)) \xi(T) | \xi(t), Y(t)]}{\xi(t)} \equiv F(\xi(t), Y(t), t), \end{aligned} \quad (65)$$

where we recall that $f(y) \equiv \inf\{x \geq 0 : V'(x) \leq y\}$ and where the third equality follows from the aforementioned fact that (ξ, Y) is a diffusion process.

The rest is now standard. We will use the martingale connection of $F(\xi, Y, t)$ under Q to compute the replicating strategy for W_T^* and the partial differential equation F satisfies. Using Itô's lemma to express $e^{-\int_0^t r(s) ds} F(\xi(t), Y(t), t)$ under the martingale measure to conclude that the time trend term must be zero, we get a second order linear partial differential equation:

$$\frac{1}{2} F_{\xi\xi} \xi^2 |\kappa|^2 + \frac{1}{2} \text{tr}(F_{YY} g g^T) - F_Y^T \xi g \kappa^T - F_\xi [r - |\kappa|^2] + F_Y^T [\beta + g \kappa^T] - rF + F_t = 0. \quad (66)$$

In addition,

$$e^{-\int_0^t r(s) ds} F(\xi(t), Y(t), t) = F(\xi(0), Y(0), 0) + \int_0^t \theta(s)^T dS^*(s),$$

where

$$\begin{aligned}\theta(t) = & -F_\xi(\xi(t), Y(t), t)\xi(t)I_{S^{-1}}(t)(\sigma(t)\sigma(t)^\top)^{-1}(\mu(t) - r(t)) \\ & + I_{S^{-1}}(t)(g(t)\sigma(t)^{-1})^\top F_Y(\xi(t), Y(t), t),\end{aligned}\quad (67)$$

where $I_{S^{-1}}(t)$ is a $N \times N$ diagonal matrix with the n -th diagonal element equal to $S_n(t)^{-1}$. Then putting

$$\alpha(t) \equiv \frac{F(\xi(t), Y(t), t) - \theta(t)^\top S(t)}{B(t)},$$

(α, θ) is the replicating strategy for W_T^* .

As in the Black-Scholes case, we can show that F is a decreasing function of ξ . Thus if we have an analytic expression for F , we define $\xi(0) = \lambda = F^{-1}(W_0, Y(0), 0)$, where F^{-1} denotes the inverse of F in its first argument. Then the process ξ is well-defined and so is the replicating strategy. Otherwise we compute F numerically by solving the partial differential equation of (66) with the boundary conditions $\lim_{\xi \rightarrow \infty} F(\xi, Y, t) = 0$ and $\lim_{t \rightarrow T} F(\xi, Y, t) = f(\xi)$ and define $\xi(0)$ similarly.

The optimal trading strategy characterized in (67) has the $(M + 2)$ -fund separation property – the optimal portfolio is composed of a portfolio on the instantaneous portfolio frontier or an *instantaneous frontier portfolio*, M portfolios whose returns are perfectly correlated with the unanticipated changes in the state variables Y , and the riskless security. To see this, we first convert the optimal numbers of shares of the risky securities held to the optimal proportions of wealth invested in the risky securities. Let $A_n(t)$ be the optimal proportion of the wealth invested in the n -th risky security at time t and let $A(t) = (A_1(t), \dots, A_N(t))^\top$. Note that $A(t)$ is the vector of optimal portfolio weights in the risky securities. Then

$$A(t) = -\frac{F_\xi(\xi(t), Y(t), t)\xi(t)}{F(\xi(t), Y(t), t)}(\sigma(t)\sigma(t)^\top)^{-1}(\mu(t) - r(t)\mathbf{1}_N) + (g(t)\sigma(t)^{-1})^\top \frac{F_Y(\xi(t), Y(t), t)}{F(\xi(t), Y(t), t)}, \quad (68)$$

where $\mathbf{1}_N$ is an N vector of 1's. The optimal proportion invested in the riskless security at time t is $(1 - A(t)^\top \mathbf{1})$. For brevity, we will henceforth use $F(t)$ to denote $F(\xi(t), Y(t), t)$ and similarly use $F_\xi(t)$ and $F_Y(t)$ for $F_\xi(\xi(t), Y(t), t)$ and $F_Y(\xi(t), Y(t), t)$, respectively.

Note that $A(t)$ only depends on $\xi(t)$, $Y(t)$, and t . Thus being viewed as the optimal control, $A(t)$ is a feedback control by noting $\xi(t) = F^{-1}(W(t), Y(t), t)$, where $W(t) = \alpha(t)B(t) + \theta(t)^\top S(t)$.

Second, we recall from Huang (1990b, relation (12)) that $(\sigma(t)\sigma(t)^\top)^{-1}(\mu(t) - r(t))$ is proportional to the portfolio weights on the risky securities of a frontier portfolio at time t , where the portfolio frontier is generated using the instantaneous expected rates of return on securities and the variance-covariance matrix of the instantaneous returns, all at time t . Thus the first term on the right-hand side of (68) is proportional to the portfolio weights on risky securities of a *instantaneous frontier portfolio*.

Third, each column of $(g(t)\sigma(t)^{-1})^\top$ can be viewed as the weights on the risky securities of a portfolio. Thus $(g(t)\sigma(t)^{-1})^\top$ represents the portfolio weights on the risky securities of M portfolios. Note that

$$(g(t)\sigma(t)^{-1})\sigma(t)g(t)^\top = g(t)g(t)^\top;$$

that is, the covariance matrix at time t of the returns on these M portfolios with the changes of Y is equal to the variance-covariance matrix of Y . This implies that the m -th column of $(g(t)\sigma(t)^{-1})^\top$

is proportional to the weights on the risky securities of a portfolio whose returns at time t are perfectly correlated with the changes in Y .

In summary, the optimal portfolio for an investor is composed of $M + 2$ parts: the riskless security, an instantaneous frontier portfolio, and M portfolios having returns perfectly correlated with the changes in Y . The proportions of wealth invested in these $M + 2$ parts are determined by the investor's utility function as reflected in the function F and its derivatives.

The implication for this mutual fund separation result is profound. If we can identify the state variables Y , then we only need the riskless security and $M + 1$ mutual funds – one instantaneous frontier portfolio and M portfolios whose returns perfectly correlated with the changes in Y for all investors to trade to their optimal portfolios. When the number M is much smaller than N , the number of mutual funds needed is much smaller than the number of securities available.

3 The Intertemporal Capital Asset Pricing Model

The previous section characterizes the optimal portfolio held by an investor. In this section, we will demonstrate the implications of the optimality condition on the equilibrium relation among the expected instantaneous rates of return on securities. For brevity, we will simply say the rates of return on the securities with the understanding that these returns are always instantaneous returns.

For simplicity we assume that there exists a single representative investor in the economy and the riskless security is in zero supply. Then this investor's wealth over time $W(t)$ is the aggregate wealth in the economy and his optimal portfolio weights on the risky securities $A(t)$ is the *market portfolio*. Note that since the riskless security is in zero supply, all the aggregate wealth must be invested in the risky securities in the market equilibrium. Thus $A(t)^\top \mathbf{1}_N = 1$ for all t and the rate of return on the market portfolio at time t is

$$A(t)^\top dS(t)/S(t) = A(t)^\top \mu(t)dt + A(t)^\top \sigma(t)dw(t).$$

Premultiplying (68) by $(\sigma(t)\sigma(t)^\top)^{-1}$ we get

$$\mathbf{V}_{S,W}(t) \equiv (\sigma(t)\sigma(t)^\top)A(t) = \frac{-F_\xi(t)}{F(t)}[\mu(t) - r(t)\mathbf{1}_N] + \mathbf{V}_{S,Y}(t)\frac{F_Y(t)}{F(t)},$$

where $\mathbf{V}_{S,W}(t)$ is the covariance matrix between the returns on the N risky securities and that on the market portfolio at time t and $\mathbf{V}_{S,Y}(t) = \sigma(t)g(t)^\top$ is the covariance matrix between the returns on the N risky securities and the unanticipated changes in Y at time t . Rearranging the above relation we get

$$\begin{aligned} \mu(t) - r(t)\mathbf{1}_N &= \frac{-F(t)}{F_\xi(t)}\mathbf{V}_{S,W}(t) + \mathbf{V}_{S,Y}(t)\frac{F_Y(t)}{F_\xi(t)} \\ &= \mathbf{V}_{S,WY}(t) \begin{pmatrix} -F(t)/F_\xi(t) \\ F_Y(t)/F_\xi(t) \end{pmatrix}, \end{aligned} \tag{69}$$

where

$$\mathbf{V}_{S,WY}(t) = (\mathbf{V}_{S,W}(t), \mathbf{V}_{S,Y}(t)).$$

That is, the expected rates of return on risky securities are linearly related to the covariances between their returns with that of the market portfolio and with the unanticipated changes in the

state variables Y . The weights on these covariances depend on F and its derivatives, which in turn depend on the investor's utility function.

We now proceed to show that we can substitute observables for $F(t)/F_\xi(t)$ and $F_Y(t)/F_\xi(t)$ in (69). First, premultiply (69) by $A(t)^\top$ to get

$$\mu_W(t) - r(t) \equiv A(t)^\top [\mu(t) - r(t)\mathbf{1}_N] = \mathbf{V}_{W,WY}(t) \begin{pmatrix} -F(t)/F_\xi(t) \\ F_Y(t)/F_\xi(t) \end{pmatrix}, \quad (70)$$

where $\mu_W(t)$ is the expected rate of return on the market portfolio at time t and $\mathbf{V}_{W,WY}(t) \equiv A(t)^\top \mathbf{V}_{S,WY}(t)$. Second, premultiply (69) by $g(t)\sigma(t)^{-1}$ to get

$$\mu_Y(t) - r(t)\mathbf{1}_M \equiv g(t)\sigma(t)^{-1} [\mu(t) - r(t)\mathbf{1}_N] = \mathbf{V}_{Y,WY}(t) \begin{pmatrix} -F(t)/F_\xi(t) \\ F_Y(t)/F_\xi(t) \end{pmatrix}, \quad (71)$$

where $\mu_Y(t)$ is the vector of expected rates of return on the M portfolios whose returns are perfectly correlated with the unanticipated changes of Y at time t and $\mathbf{1}_M$ is an M vector of 1's.

Combining (70) and (71) and solving for $F(t)/F_\xi(t)$ and $F_Y(t)/F_\xi(t)$ gives

$$\begin{pmatrix} -F(t)/F_\xi(t) \\ F_Y(t)/F_\xi(t) \end{pmatrix} = \mathbf{V}_{WY}(t)^{-1} \begin{pmatrix} \mu_W(t) - r(t) \\ \mu_Y(t) - r(t)\mathbf{1}_M \end{pmatrix}.$$

Substituting this relation into (69) gives

$$\begin{aligned} \mu(t) - r(t)\mathbf{1}_N &= \mathbf{V}_{S,WY}(t) \mathbf{V}_{WY,WY}(t)^{-1} \begin{pmatrix} \mu_W(t) - r(t) \\ \mu_Y(t) - r(t)\mathbf{1}_M \end{pmatrix} \\ &\equiv \beta_{S,WY}(t) \begin{pmatrix} \mu_W(t) - r(t) \\ \mu_Y(t) - r(t)\mathbf{1}_M \end{pmatrix}. \end{aligned}$$

This is the *Intertemporal Capital Asset Pricing Model* of Merton (1973). The expected rates of return on the risky securities are linearly related to their "beta's" with respect to the returns on the market portfolio and on the M portfolios whose returns are perfectly correlated with the unanticipated changes in the state variables. These beta's are the multiple regression coefficients of the rates of return on the risky securities on those on the market portfolio and the M portfolios perfectly correlated with Y . The weights on these beta's are the risk premiums on the market portfolio and on the M portfolios perfectly correlated with Y .

The Intertemporal Capital Asset Pricing Model is the generalization of the Capital Asset Pricing Model of Sharpe (1964), Lintner (1965), and Mossin (1966). There the equilibrium expected rates of return on the securities are linearly related to their beta's with respect to the market portfolio. Here, since the return distributions change over time and the changes are determined by the changes in the state variables Y , besides the beta's with respect to the market portfolio, M more beta's arise. These M beta's capture the sensitivity of security returns to the changes in the state variables.

4 Concluding Remarks

We have summarized how the modern theory of contingent security valuation can be useful in deriving the optimal consumption and portfolio policy for an investor. For simplicity, we have

focused our discussion on the case where an investor only consumes at a single date T while trading in securities are allowed any time before T . The generalization to the cases where consumption before time T is allowed can be easily accomplished. The reader is referred to Cox and Huang (1986, 1989) for details.

Our hypothesis about the price processes guarantees that the markets are dynamically complete. This is more restrictive than the traditional dynamic programming approach as briefly reviewed in Huang (1990a, section 3). Our technique can be generalized to situations where the markets are not dynamically complete. The reader is referred to He and Pearson (1989) and Pagès (1989) for details.

We derived the *Intertemporal Capital Asset Pricing Model* of Merton (1973). When there is intermediate consumption, similar techniques can be used to derive the *Intertemporal Consumption Capital Asset Pricing Model* of Breeden (1979) discussed in Huang (1990a, section 3). The reader is referred to Huang (1989) for details.

Throughout our discussion of the intertemporal capital asset pricing model, we assume that an equilibrium exists with prices as posited. For the existence of an equilibrium and related issues see Duffie (1986), Duffie and Huang (1985), and Huang (1985a, 1985b, 1987).

5 References

1. Breeden, D. 1979. An intertemporal capital pricing model with stochastic investment opportunities. *Journal of Financial Economics* 7: 265–296.
2. Cox, J., and C. Huang. 1986. A variational problem arising in financial economics. to appear in *Journal of Mathematical Economics*.
3. Cox, J., and C. Huang. 1989. Optimal consumption and portfolio policies when security prices follow a diffusion process. *Journal of Economic Theory* 49:33–83.
4. Cox, J., J. Ingersoll, and S. Ross. 1985. An intertemporal general equilibrium model of security prices. *Econometrica* 53:363–384.
5. Duffie, D., and C. Huang. 1985. Implementing Arrow–Debreu equilibria by continuous trading of few long-lived securities. *Econometrica* 53:1337–1356.
6. Duffie, D. 1986. Stochastic equilibria: Existence, spanning number, and the “no expected gain from trade” hypothesis. *Econometrica* 54:1161–1184.
7. He, H., and N. Pearson. 1989. Consumption portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. Mimeo. To appear in *Journal of Economic Theory*. Technology.
8. Huang, C. 1985a. Information structure and equilibrium security prices. *Journal of Economic Theory* 35:33–71.
9. Huang, C. 1985b. Information structure and viable price systems. *Journal of Mathematical Economics* 13:215–240.

10. Huang, C. 1987. An intertemporal general equilibrium security pricing model: The case of diffusion information. *Econometrica* 55:117-142.
11. Huang, C., 1989, *Lecture Notes on Advanced Financial Economics*, Sloan School of Management, Massachusetts Institute of Technology.
12. Huang, C., 1990a. An overview of modern financial economics. The first article in this collection.
13. Huang, C., 1990b. The static portfolio theory and the Capital Asset Pricing Model. The second article in this collection.
14. Huang, C., 1990c. Contingent securities valuation in continuous time. The third article in this collection.
15. Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47:13-37.
16. Merton, R. 1971. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory* 3:373-413.
17. Merton, R. 1973a. An intertemporal capital asset pricing model. *Econometrica* 41:867-887.
18. Mossin, J. 1966. Equilibrium in a capital asset market. *Econometrica* 35:768-783.
19. Pagès, H. 1989. Three Essays In Optimal Consumption. Unpublished Ph.D. thesis. Massachusetts Institutes of Technology.
20. Sharpe, W. 1964. Capital asset prices: A theory of capital market equilibrium under conditions of risk. *Journal of Finance* 19:425-442.

The Term Structure of Interest Rates and the Pricing of Interest Rate Sensitive Securities

Chi-fu Huang

1 Introduction

In this final piece of the series of reviews of modern financial economics, we will apply the theory of contingent securities valuation discussed in Huang (1990a) to study the pricing of default-free bonds, or the so-called *term structure of interest rates*. We will also mention how the knowledge of the term structure of interest rates can help us understand the pricing of other interest rate sensitive securities such as options and futures contracts on bonds.

2 Arbitrage Pricing of Default-Free Discount Bonds

We will focus our attention here in the section on the pricing of default-free discount bonds. A default-free unit discount bond is a financial security that pays \$1 for sure at a fixed future date, which is termed the *maturity date*. We assume that discount bond prices of all maturities are functions of an N -dimensional “state variables” Y , and let $S(Y(t), t; T)$ denote the price at time t of a unit discount bond with a maturity date T . These state variables follow a diffusion process:

$$Y(t) = Y(0) + \int_0^t \beta(Y(s), s) ds + \int_0^t g(Y(s), s) dw(s), \quad \forall t,$$

where w is an N dimensional standard Brownian motion under a probability P , β is an $N \times 1$ vector, and g is an $N \times N$ matrix. We assume that g is of full rank for all Y and t , and S is a smooth function of $Y(t)$ and t .

Besides discount bonds, there also exists an instantaneous borrowing and lending opportunity. The instantaneous riskless interest rate at time t is $r(Y(t), t)$. We put $B(t) = \exp\{\int_0^t r(Y(s), s) ds\}$ and call this the price process for the riskless asset. For simplicity of notation, we will often denote $S(Y(t), t; T)$ and $r(Y(t), t; T)$ by $S(t; T)$ and $r(t)$, respectively.

Consider now a collection of default-free unit discount bonds with maturities $T_1, T_2, \dots, T_m$, where $m > N$. Without loss of generality, assume that $T_1 < T_2 < \dots < T_m$. In addition, we will impose that the prices of these discount bonds admit no arbitrage opportunities.

We know from the discussions in Huang (1990a, 1990b) that a necessary condition for there to be no arbitrage opportunities is the existence of an equivalent martingale measure. That is, there

exists a probability measure Q under which $S(t; T_i)/B(t)$ is a martingale for all $i = 1, 2, \dots, m$. Using Itô's lemma we can write

$$\begin{aligned} d(S(t; T_i)/B(t)) = & [-S(t; T_i)r(t)/B(t) + (\mathcal{L}S(t; T_i))/B(t) + S_t(t; T_i)/B(t)]dt \\ & + S_y(t; T_i)^\top g(Y(t), t)/B(t)dw(t), \end{aligned} \quad (72)$$

where

$$\mathcal{L}S(t; T_i) = \frac{1}{2} \text{tr}(S_{yy}(t; T_i)g(Y(t), t)g(Y(t), t)^\top) + S_y(t; T_i)^\top \beta(Y(t), t) + S_t(t; T_i),$$

tr denotes the “trace”, and the subscripts of S are partial derivatives of S . Note that $(\mathcal{L}S(t; T_i) + S_t(t; T_i))/S(t; T_i)$ is the instantaneous expected rate of return at time t on the discount bond with maturity T_i . From Girsanov theorem, there exists an N -dimensional process $\kappa(t)$ so that under Q ,

$$w^*(t) = w(t) - \int_0^t \kappa(s)ds \quad t \in \mathbb{R}_+, \quad (73)$$

is a standard Brownian motion; see Huang (1990a, theorem 1). Substituting (73) into (72) gives

$$\begin{aligned} d(S(t; T_i)/B(t)) = & [-S(t; T_i)r(t)/B(t) + (\mathcal{L}S(t; T_i))/B(t) + S_t(t; T_i)/B(t) \\ & + S_y(t; T_i)^\top g(Y(t), t)\kappa(t)/B(t)]dt \\ & + S_y(t; T_i)^\top g(Y(t), t)/B(t)dw^*(t). \end{aligned} \quad (74)$$

Since $S(t; T_i)/B(t)$ is a martingale under Q and thus cannot have any time trend, we must have

$$-S(t; T_i)r(t) + \mathcal{L}S(t; T_i) + S_t(t; T_i) + S_y(t; T_i)^\top g(Y(t), t)\kappa(t) = 0, \quad \forall t. \quad (75)$$

We therefore conclude that a necessary condition for there not to have any arbitrage opportunity is the existence of an N -dimensional process κ so that (75) holds for all i . Given that the number of discount bonds with different maturities is strictly greater than the dimension of the Brownian motion, that is, $M > N$, elementary linear algebra shows that κ must be uniquely determined. Furthermore, since the maturities are arbitrarily chosen, κ must be independent of the maturities and thus be functions of only Y and t . We therefore conclude that $\kappa(t) = \kappa(Y(t), t)$.

Now note the following. If we knew what $\kappa(Y(t), t)$ is, then (75) would be a partial differential equation that $S(Y, t; T)$ must satisfy:

$$-S(Y, t; T_i)r(Y, t) + \mathcal{L}S(Y, t; T_i) + S_t(Y, t; T_i) + S_y(Y, t; T_i)^\top g(Y, t)\kappa(Y, t) = 0, \quad \forall t \in [0, T_i)Y. \quad (76)$$

In addition, the value at the maturity date of a default-free unit discount bond must be equal to 1. Thus $S(Y, t; T_i)$ is a solution to (76) with a boundary condition $S(Y, T_i; T_i) = 1$.

In summary, the arbitrage approach of the pricing of default-free discount bonds lies in the specification of the function $\kappa(Y, t)$. This specification cannot be arbitrary since the function κ must be chosen so that the condition of Girsanov theorem is satisfied. That is, κ must be such that $E[\zeta(T)] = 1$ for all $T > 0$, where

$$\zeta(T) = \exp \left\{ -\frac{1}{2} \int_0^T \kappa(Y(s), s)^\top \kappa(Y(s), s)ds + \int_0^T \kappa(Y(s), s)^\top dw(s) \right\}.$$

Moreover, once a $\kappa(Y, t)$ is chosen to satisfy the above condition, we know that

$$Q(A) = \int_A \zeta(\omega, T) P(d\omega)$$

for all events A distinguishable from observing the state variables Y , is a probability equivalent to P . Then the discount bond prices given by

$$S(Y(t), t; T_i) = B(t) E_t^*[1/B(T_i)] = E^*[\exp\{-\int_t^{T_i} r(Y(s), s) ds\} | Y(t)],$$

where $E_t^*[\cdot]$ is the conditional expectation under Q at time t and where the second equality follows from the Markov property of Y . It is then easily seen that $S(t; T_i)/B(t)$ is a martingale under Q with $S(T_i; T_i) = 1$. Discount bond prices for all maturities so defined thus form an arbitrage-free price system. Of course, $S(Y, t; T_i)$ will satisfy (76) with the boundary condition $S(Y, T_i; T_i) = 1$ (provided that S is a smooth function of Y and t).

Given the unit discount bond prices, the yield-to-maturity of the i -th bond at time $t < T_i$ is defined to be the number $R(Y, t; T_i)$ such that

$$S(Y, t; T_i) = \exp\{-R_i(Y, t; T_i)(T_i - t)\}.$$

Hence, $R(Y, t; T_i) = \frac{-\ln S(Y, t; T_i)}{T_i - t}$. The *yield curve* at time t , or the term structure of interest rates at time t , is just the function $R(Y, t; T_i)$ for various different T_i when the state variables are Y .

Before leaving this section, we note that since coupon bonds are nothing more than portfolios of discount bonds, once we know how to value discount bonds we know how to value coupon bonds.

3 The Cox-Ingersoll-Ross Model

The Cox, Ingersoll, Ross (1985) (CIR) model of the term structure of interest rates is a special case of the arbitrage approach outlined in the previous section. Assume that the dimension of Y is one and the single state variable is the instantaneous interest rate $r(t)$. Assume further that r follows a mean reverting process:

$$r(t) = r(0) + \int_0^t k(\theta - r(s)) ds + \int_0^t \sigma \sqrt{r(s)} dw(s), \quad (77)$$

where w is a one dimensional standard Brownian motion, and $k > 0$, $\theta > 0$ and $\sigma > 0$ are the three parameters. The parameter θ is the *long-run mean* of r . When the current realization of the interest rate is below its long-run mean, the time trend in the instant after will be positive pulling r towards its long-run mean. Similarly, if the current interest rate is above the long-run mean, the time trend in the instant after will be negative pulling r down towards its long-run mean. The parameter k represents the intensity of the mean reversion. The mean reverting time trend is also subject to random shocks governed by a Brownian motion. The dispersion of the random shock is proportional to the square-root of the interest rate with a proportionality factor σ . The higher the interest rates, the greater the uncertainty centered around the movements of the interest rates.

Using equilibrium arguments, CIR specify

$$\kappa(r, t) = -\frac{\lambda}{\sigma}\sqrt{r},$$

where λ is a scalar which may be positive or negative. The interest rate process under the martingale measure becomes

$$r(t) = r(0) + \int_0^t (k + \lambda) \left(\frac{k\theta}{k + \lambda} - r(s) \right) ds + \int_0^t \sigma \sqrt{r(s)} dw^*(s), \quad (78)$$

which is also a mean-reverting process with a long-run mean $k\theta/(k + \lambda)$ and an intensity of mean reversion $k + \lambda$. Note better that there are four parameters for the process of r under Q , while there are three parameters under P . The one additional parameter is λ .

The partial differential equation of (76) simplifies to be

$$\frac{1}{2}\sigma^2 r S_{rr}(r, t; T_i) + S_r(r, t; T_i)[k(\theta - r) - \lambda r] + S_t(r, t; T_i) - rS(r, t; T_i) = 0, \quad (79)$$

with the boundary condition $S(r, T_i; T_i) = 1$. One verifies that

$$S(r, t; T_i) = A(t, T_i)e^{-r(t)G(t, T_i)},$$

where

$$\begin{aligned} A(t, T_i) &= \left[\frac{2\gamma e^{(k+\lambda+\gamma)(T_i-t)/2}}{(\gamma + k + \lambda)(e^{\gamma(T_i-t)} - 1) + 2\gamma} \right]^{2k\theta/\sigma^2}, \\ G(t, T_i) &= \frac{2(e^{\gamma(T_i-t)} - 1)}{(\gamma + k + \lambda)(e^{\gamma(T_i-t)} - 1) + 2\gamma}, \\ \gamma &= \left[(k + \lambda)^2 + 2\sigma^2 \right]^{1/2}. \end{aligned}$$

Itô's lemma allows us to write

$$dS(r, t; T_i) = r[1 - \lambda G(t, T_i)]S(r, t; T_i)dt - G(t, T_i)S(r, t; T_i)\sigma\sqrt{r}dw(t).$$

Thus if $\lambda < (>)0$, the instantaneous expected rate of return on a discount bond that is not maturing in the next instant is greater (smaller) than r .

The yield-to-maturity at time t of a discount bond maturing at T_i is

$$R(r, t; T_i) = [rG(t, T_i) - \ln A(t, T_i)]/(T_i - t).$$

As maturity nears, $R(r, t; T_i)$ approaches r . As we consider longer and longer maturities, $R(r, t; T_i)$ approaches a limit:

$$\lim_{T_i \rightarrow \infty} R(r, t; T_i) = \frac{2k\theta}{\gamma + k + \lambda}.$$

One can also show that when $r < \theta$, the yield curve is uniformly rising. With r in excess of $k\theta/(k + \lambda)$, the yield curve is uniformly falling. For intermediate values of the interest rates, the yield curve is humped.

At first sight, the CIR model is quite restrictive as, for example, it cannot produce a yield curve that has two humps. As a consequence, it does not have the feature that there always exists a set of parameters so that an empirically observed yield curve is consistent with the model. This “shortcoming” however can be easily overcome by allowing some of the parameters to be time dependent. For example, we can let the parameter θ to be a function of time. The instantaneous interest rate process becomes

$$r(t) = r(0) + \int_0^t k(\theta(s) - r(s))ds + \int_0^t \sigma \sqrt{r(s)}dw(s). \quad (80)$$

Then the discount bond price at time t is

$$S(r, t; T_i) = \hat{A}(t, T_i)e^{-G(t, T_i)r},$$

where

$$\hat{A}(t, T_i) = \exp \left(-k \int_t^{T_i} \theta(s)G(s, T_i)ds \right).$$

Then the shape of the yield curve can take any arbitrary shape depending on the function $\theta(t)$.

If the model is correct, one should be able to infer the values of the parameters from the observed default-free bond prices.

4 Pricing of Interest Rate Sensitive Securities

A model of the term structure of interest rates is not only useful for the valuation of default-free bonds. It is also useful for the valuation of all the interest rate sensitive securities. We assume throughout that CIR model is correct and we know the parameters k , θ , σ , and λ .

Consider a European call option on a default-free discount bond. The maturity date of the bond is T_i . The maturity date and the exercise price of the call option are $T < T_i$ and $K > 0$, respectively. Let the price of the option at time t be $C(t)$. Since there is only one Brownian motion that drives the uncertain interest rate and there are many bonds with different maturities, the markets are dynamically complete. The option can thus be priced by arbitrage. By the martingale property of arbitrage-free prices, we know

$$\begin{aligned} C(t) &= B(t)E_t^*[\max[S(r(T), T; T_i) - K, 0]/B(T)] \\ &= E^* \left[\exp \left\{ - \int_t^T r(s)ds \right\} \max[S(r(T), T; T_i) - K, 0] | r(t) \right], \end{aligned}$$

where the second equality follows from the Markov property of r under Q . Since the transition probabilities of r under the equivalent martingale measure Q is known (see CIR, p.391), the above expectation can be explicitly evaluated. The interested reader can consult CIR (p.396) for a closed-form expression.

Applying the same principle we can value any interest rate sensitive security. Let $\phi(t)$ be the random cash flow of a security at date t whose value depends on the history of r from time 0 to

time t . Assume for simplicity that $\phi(t)$ is zero except at a finite number of dates $t_1, t_2, \dots, t_n$. The value at t of $\phi(t_i)$ with $t_i > t$ is

$$E_t^* \left[e^{-\int_t^{t_i} r(s) ds} \phi(t_i) \right].$$

Thus the value at t of the security is

$$\sum_{t_i > t} E_t^* \left[e^{-\int_t^{t_i} r(s) ds} \phi(t_i) \right].$$

An example of a general interest rate sensitive security is a pool of home mortgages or *Mortgage-Backed Securities*. In this example, $\phi(t_i)$ is the cash flow at time t_i generated by interest rate payments and principal prepayments of the mortgages in the pool. To value such a security, one only need to have two things: The parameters of the instantaneous default-free interest rate process under Q and the projections of future cash flows of the pool of mortgages.

5 Futures and Forward Prices on Interest Rates Sensitive Securities

We will now discuss how to value continuously resettled financial contracts such as forward and futures contracts on discount bonds. For concreteness of our discussion, we continue to use the CIR model of interest rates.

An individual who takes a long position in a forward contract agrees to buy a designated good or security on a specified future date, the *maturity date*, for the *forward price* prevailing at the time the contract is initiated. At the maturity date, the forward price is equal to the spot price for the good or security.

Formally, consider a discount bond with a maturity date T_i . Denote the forward price at time t for this bond for the delivery date $T < T_i$ by $F(t)$. Since no money changes hand when a forward contract is initiated, $F(t)$ must be determined so that the contract is worthless at time t . That is, the value at t of a payoff at time T of $S(r(T), T; T_i) - F(t)$ is zero. Mathematically, $F(t)$ must be such that

$$E_t^* [e^{-\int_t^T r(s) ds} (S(r(T), T; T_i) - F(t))] = 0.$$

Note first that as $S(t; T_i)/B(t)$ is a martingale under Q , $E_t^* [e^{-\int_t^T r(s) ds} S(r(T), T; T_i)] = S(r(t), t; T_i)$. Next note that as $F(t)$ is determined at time t , we have

$$E_t^* [e^{-\int_t^T r(s) ds} F(t)] = F(t) E_t^* [e^{-\int_t^T r(s) ds}] = F(t) S(r(t), t; T),$$

where the second equality follows from the fact that $S(r(T), T; T) = 1$ since it is a discount bond maturing at time T . These imply that

$$F(t) = S(r(t), t; T_i) / S(r(t), t; T). \quad (81)$$

We have thus expressed the forward price at time t for a discount bond by the then prevailing bond prices. It is clear from the expression of (81) that, besides the parameters T and T_i , $F(t)$ is a function of $r(t)$ and t .

Now we turn our attention to futures prices. An individual who is long in a futures contract nominally agrees to buy a designated good or security on the maturity date, for the *futures prices* prevailing at the time the contract is initiated. At maturity date, futures price is equal to the spot price. No money changes hands initially. Subsequently, however, as the futures price changes, the party in whose favor the price change occurred, must immediately be paid the full amount of the change by the losing party. As a result, the payment required on the maturity date to buy the underlying good or security is simply its spot price at that time. The difference between that amount and the initial futures price has been paid (or received) in installments throughout the life of the contract. Futures prices must change continually over time in such a fashion that the remaining stream of future payments described above always has a value of zero.

Let $f(r, t)$ be the futures price at time t for a unit discount bond delivered at T . The maturity date of the underlying discount bond is $T_i > T$. Assume for now that $f(r, t)$ is a twice continuously differentiable function of $r(t)$ and once continuously differentiable function of t .

By definition, the cash flow due to the holder of a futures contract over the interval $[s, s + ds]$ is the change of the futures price over the interval $df(r(s), s)$. By the definition of futures prices, the present value at time t of all these cash flows for all $s \in [t, T]$ must be equal to zero. That is,

$$E_t^* \left[\int_t^T e^{-\int_t^s r(\tau) d\tau} df(r(s), s) \right] = 0.$$

As f is a continuous function and a realization over time of the interest rate is a continuous function of time, integration by parts gives

$$f(r(t), t) = E_t^* \left[e^{-\int_t^T r(\tau) d\tau} f(r(T), T) + \int_t^T e^{-\int_t^s r(\tau) d\tau} r(s) f(r(s), s) ds \right]. \quad (82)$$

Multiplying both sides by $e^{-\int_0^t r(\tau) d\tau}$ and then adding $\int_0^t e^{-\int_0^s r(\tau) d\tau} r(s) f(r(s), s) ds$ to both sides give

$$\begin{aligned} & e^{-\int_0^t r(\tau) d\tau} f(r(t), t) + \int_0^t e^{-\int_0^s r(\tau) d\tau} r(s) f(r(s), s) ds \\ &= E_t^* \left[e^{-\int_0^T r(\tau) d\tau} f(r(T), T) + \int_0^T e^{-\int_0^s r(\tau) d\tau} r(s) f(r(s), s) ds \right]. \end{aligned} \quad (83)$$

The right-hand-side of the above relation is a conditional expectation of a fixed random variable and thus is a martingale under Q . Thus the left-hand-side is also a martingale under Q . Now applying Itô's lemma to the left-hand-side and using the fact that it is a martingale under Q and cannot have any time trend we get

$$\frac{1}{2} f_{rr}(r, t) r \sigma^2 + f_r(r, t) [k(\theta - r) - \lambda r] + f_t = 0. \quad (84)$$

This is a partial differential equation that f satisfies. The boundary condition comes from the fact that at the maturity date of the futures contract and the futures price must be equal to the spot price. Hence $f(r, T) = S(r, T; T_i)$. One easily verifies that the solution to (84) with the aforespecified boundary condition is

$$f(r(t), t) = E^*[S(r(T), T; T_i) | r(t)]. \quad (85)$$

Thus the futures price process is a martingale under the martingale measure. This is a departure from the martingale result for long-lived securities, where the price processes *in units of the riskless security* is a martingale under Q . This is so because the futures price is *not* the price of a long-lived security. It is the price for a continuously resettled financial contract.

Note that since we know the transition density function of r under Q , the expectation of (85) can be explicitly calculated, which we leave for the interested reader. This explicit analytic expression is a twice continuously differentiable function of r and once continuously differentiable function of t , as we assumed to begin with.

6 Concluding Remarks

In this paper we have briefly discussed how the theory of arbitrage pricing of contingent securities can be applied to study the behavior of interest rates and the valuation of interest rate sensitive securities. Much of the discussion is taken from Huang (1989).

Besides the Cox, Ingersoll, Ross (1985) model of term structure of interest rates, the interested reader should also consult Brennan and Schwartz (1979), Heath, Jarrow, and Morton (1988), Ho and Lee (1986), Richard (1978), and Vasicek (1977).

7 References

1. Cox, J., J. Ingersoll, and S. Ross. 1985. A theory of the term structure of interest rates, *Econometrica* 53:385-407.
2. Brennan, M., and E. Schwartz. 1979. A continuous time approach to the pricing of bonds. *Journal of Banking and Finance* 3:133-155.
3. Heath, D., R. Jarrow, and A. Morton. 1988. Bond Pricing and the term structure of interest rates: a new methodology. Unpublished manuscript, Cornell University.
4. Ho, T., and S. Lee. 1986. Term structure movements and pricing interest rate contingent claims. *Journal of Finance* 41:1011-1028.
5. Huang, C., 1989, *Lecture Notes on Advanced Financial Economics*, Sloan School of Management, Massachusetts Institute of Technology.
6. Huang, C., 1990a. Contingent securities valuation in continuous time. The third article in this collection.
7. Huang, C. 1990b. Dynamic Portfolio Theory and the Intertemporal Capital Asset Pricing Model. The fourth article in this collection.
8. Richard, S. 1978. An arbitrage model of the term structure of interest rates. *Journal of Financial Economics* 6:33-57.
9. Vasicek, O. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5:177-188.

Date Due

OCT 15 1991	
13 1992	
AUG 11 1992	
MAY 21 1993	
AUG 11 1993	
APR 21 1994	
MAR 25 1997	
OCT. 24 1997	

MIT LIBRARIES DUPL



3 9080 00686241 8

